POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Path-Dependent Volatility models: how fast is SPX and VIX pricing?

Tesi di Laurea Magistrale in
Mathematical Engineering - Ingegneria Matematica

Author: **Davide Cestaro**

Student ID: 10616107
Advisor: Prof. Michele Azzone
Academic Year: 2022-23

Abstract

Over the years, much more attention has been paid to modeling asset pricing dynamics than volatility dynamics. In the recent period, however, the study of volatility trends within pricing models has increased. It has been shown that volatility is, for most of its value, path dependent; in fact, it is possible to explain up to 90% of the variance of the implied volatility of equity indices endogenously, that is, through the values of past returns alone, without the need to introduce a stochastic component such as Brownian motion. The growing interest in volatility dynamics is also given by the fact that in finance we want to do the joint smile calibration of the S&P 500/VIX implied volatility. Thus, a new model is introduced that is capable of doing this calibration and in which volatility is defined in such a way as to satisfy the "Zumbach Effect" and the "Leverage Effect."

Having defined the model, it is crucial to analyze the pricing techniques in order to produce truly applicable results. In order to be able to represent the implied volatility curve of both the S&P 500 and the VIX, one must start by calculating the derivatives (European options). Two techniques will be used: Monte Carlo simulations and binomial tree. The Monte Carlo method is well known to the world of finance and easily applicable, but in the presence of an index with a complex definition such as the VIX it does not turn out to be the only valid solution. Thus, the non-recombinant binomial tree is developed specifically from the newly introduced model: it will turn out to be a valid alternative. We will compare the results produced by the two techniques, with a focus on computational time and costs referring to memory.

Keywords: Volatility models, Monte Carlo, Binomial Tree, SPX implied volatility, VIX implied volatility.

Abstract in lingua italiana

Nel corso degli anni si è prestata molta più attenzione alla modellazione della dinamica dei prezzi degli asset che a quella della volatilità. Negli ultimi tempi, tuttavia è aumentato lo studio dell'andamento della volatilità all'interno dei modelli di pricing. È stato dimostrato che la volatilità è, per la maggior parte del suo valore, dipendente dal percorso; infatti, è possibile spiegare fino al 90% della varianza della volatilità implicita degli indici azionari in modo endogeno, cioè attraverso i soli valori dei rendimenti passati, senza la necessità di introdurre una componente stocastica come il moto browniano. Il crescente interesse per la dinamica della volatilità è dato anche dal fatto che in finanza si vuole effettuare la calibrazione congiunta dello smile della volatilità implicita dell'S&P 500/VIX. Viene quindi introdotto un nuovo modello in grado di effettuare tale calibrazione e in cui la volatilità è definita in modo tale da soddisfare l'"Effetto Zumbach" e l'"Effetto Leverage".

Definito il modello, è fondamentale analizzare le tecniche di pricing per produrre risultati realmente applicabili. Per poter rappresentare la curva di volatilità implicita sia dell'S&P 500 che del VIX, è necessario partire dal calcolo dei derivati (opzioni europee). Verranno utilizzate due tecniche: simulazioni Monte Carlo e albero binomiale. Il metodo Monte Carlo è ben noto al mondo della finanza e facilmente applicabile, ma in presenza di un indice con una definizione complessa come il VIX non risulta essere l'unica soluzione efficace. Pertanto, l'albero binomiale non ricombinante è stato sviluppato appositamente a partire dal modello appena introdotto: si rivelerà una valida alternativa. Confronteremo i risultati prodotti dalle due tecniche, con particolare attenzione ai tempi di calcolo e ai costi relativi alla memoria.

Parole chiave: Modelli di volatilità, Monte Carlo, Albero Binomiale, Volatilità implicita SPX, Volatilità implicita VIX.

Contents

Introduction

In the financial industry, when it comes to pricing, hedging, and risk management for derivatives portfolios, a crucial aspect is understanding and effectively capturing the simultaneous movements of the underlying assets and their implied volatilities. In addition, it turns out to be very useful to develop accurate predictors to forecast future realized volatility. Recently, literature has mainly focused on examining the price dynamics of a given asset, focusing on models in which volatility exhibits a unique dynamic. Among the models explored in this context we find stochastic volatility models (SV). These models introduce a random component into the asset's dispersion indicator, which may or may not be correlated with the asset's price dynamics (introduced for the first time by Hull and White in [27]).

Through empirical analysis of the data, it was observed that implied volatility and future realized volatility depend on the path traced by the underlying asset price in the recent past; thus, the idea of introducing a model for path-dependent volatility was born: Path-Dependent Volatility model (PDV). Guyon, wanting to explain the empirically observed features, present his PDV model in [24] (2014). This model was expanded, and in 2022 a new paper [25] was published by Guyon and Lekeufack. In particular, they want to present a model in which the Leverage Effect, tendency of volatility to show a negative correlation with asset price returns, and the Zumbach Effect, relationship between past returns and future volatility, are observed. In addition, a third key factor drove Guyon and Lekeufack to present their model: the ability to replicate the joint calibration of S&P 500 and VIX implied volatility smiles. Stochastic volatility models in fact are unable to present a model that can accurately replicate the implied volatility curves of an underlying and its volatility; not even the Heston model, considered the most versatile and modern SV model, succeeds.

Guyon and Lekeufack's model uses a simple approach to path-dependency, incorporating a linear combination of the weighted sums of past daily returns and the square root of the weighted sums of past daily returns squared. Different weights are assigned to different time instants, in order to capture both short and long memory. This simple but effective model has shown greater accuracy in replicating implied and realized volatility than clas-

sical SV models. The proposed Path-Dependent Volatility model operates in continuous time and can be provided with historical or risk-neutral parameters. The weights can be approximated using exponential kernel superpositions, which represent functions of stochastic discount factors. This makes it possible to create Markovian models, in which the future state of the system depends solely on its current state, independent of past states, a crucial aspect for the application of major pricing techniques.

In this Thesis an alternative pricing technique to the widely used Monte Carlo method is proposed: the binomial tree. Thus having these two methods available we could compare their results. We will analyze the first method in different ways, implementing two possible time discretizations for the simulation, also introducing techniques to reduce variance (i.e. the error) so that the best can be obtained from this methodology. With respect to the second method, meanwhile, the entire structure for the dynamics present in the introduced model will be implemented, referring to only one temporal discretization. We will apply these procedures in calculating implied volatility of the S&P 500, comparing results, computational time and cost in terms of memory, and pointing out weaknesses and strengths.

Before presenting the results regarding the implied volatility of the VIX, an analysis will first be performed with respect to this index. Indeed, the definition will be presented, explaining the complexity of simulation and the consequent ineffectiveness of models to date in explaining its implied volatility curve. Again, in Monte Carlo method and binomial (non-recombinant) tree will be used to simulate the VIX.

Finally, the results will be presented, determining in which situations it is more appropriate to use one method or the other, considering computational time as the main term of comparison. The capability of PDV's model for joint calibration of S&P 500 and VIX implied volatility smiles will then be analyzed, presenting some ideas for how both the model and the pricing techniques used can be improved and implemented more, in order to obtain more accurate data.

Listed below the structure of this Thesis.

In Chapter 1 we go over the main models introduced in quantitative finance, paying particular attention to how volatility has been considered. We started by considering constant volatility and then introduced local volatility models, models in which volatility becomes a function of time. Finally, we analyzed stochastic volatility models, pointing out the weaknesses, obvious by the large difference in the results from empirical data.

In Chapter 2 we reconstructed and analyzed the model introduced by Guyon and Leke-

ufack, emphasizing the insights behind Path-Dependent Volatility models (PDV). We traced the construction of the model by paying special attention to the *"2-factor Markovian PDV model"* and the *"4-factor Markovian PDV model"*.

In Chapter 3 we presented the techniques used for model simulation (both 2-factor and 4-factor), focusing mainly on the Monte Carlo and binomial tree methods. We have shown in detail how to move from the theoretical model to practical applications. For the Monte Carlo method, we have introduced systems to reduce the variance of the results; while for the method via the binomial tree, we have shown its step-by-step construction.

In Chapter 4 we presented the results regarding the dynamics of the patterns, analyzing for various processes their significance. We then introduced the simulations of the S&P 500 implied volatility curves, comparing the outcomes produced by the two methods both in terms of time and accuracy. After this comparison performed with both the 2-factor and 4-factor models, we presented the results with respect to the smile of implied volatilities for the VIX.

In Chapter 5 we summarized the main results obtained throughout this Thesis, describing how for different circumstances one pricing method or the other was better. Finally, some insights were presented from which techniques can be implemented that could potentially be more effective.

1 | Quantitative models for finance

The following chapter will introduce the main quantitative models for finance, starting with the simplest ones in which volatility is seen as a constant value over time to the most complex ones in which instead volatility is endowed with a process of its own. The reasons why most advanced models do not respect the Zumbach Effect, the Leverage Effect, and the problem of joint calibration of S&P 500 and VIX option smiles, led to the introduction of the model discussed in the Thesis, will be highlighted.

1.1. Deterministic constant volatility

Advanced tools of mathematics were first introduced to the world of finance in a theoretical way in 1900 by Louis Bachelier, who in his PhD Thesis [1], hypothesized a model for the value of an asset price containing a stochastic process, which was intended to represent changes in the market. Bachelier's main purpose was to estimate a given stock by a simple sum of two factors:

$$S_t = S_0 + \sigma W_t, \tag{1.1}$$

where S_t denotes the asset price at future time instant t greater than zero (t>0), S_0 the current stock price (always grater than 0), σ the constant and deterministic value of volatility, and W_t is a Wiener process, a stochastic process such that:

- $W_0 = 0$
- $W_t \sim \mathcal{N}(0, t)$
- $W_t - W_s \perp \mathcal{F}_s$ with $t > s$
- $W_{t+h} - W_h \sim W_{s+h} - W_s \sim \mathcal{N}(0, h)$

It was immediately apparent that it was impossible to guarantee that the price of the underlying asset would always be greater than or equal to zero, so consequently years later, in 1973, Black & Scholes [4] and Merton [34] introduced their model as stochastic

differential equation:

$$\begin{cases} dS_t = rS_t dt + \sigma S_t dW_t \\ S_0 \end{cases}, \tag{1.2}$$

where the variable added r is the risk free interest rate. Using the Itô's formula[1], the Geometric Brownian Motion (GMB) can be obtained, expressed by the following equation (calculations shown in appendix A.2):

$$S_t = S_0 e^{\left(r - \frac{1}{2}\sigma^2\right)t + \sigma W_t} \tag{1.3}$$

It is evident that S_t, defined as in this equation, can only take positive values.

To explain the important aspects of this model, it is first necessary to introduce the closed formulas for the most liquid options, an European option. Given maturity T and strike K, an option is called a *"put"* if at maturity returns a payoff of $(K - S_T)^+$ and a *"call"* if at maturity returns a payoff of $(S_T - K)^+$. Using the Black and Scholes model we derive the following formulas:

$$C(S,t)_{EU} = S_t N(d_1) - Ke^{-r(T-t)} N(d_2), \tag{1.4}$$

$$P(S,t)_{EU} = Ke^{-r(T-t)} N(-d_2) - S_t N(-d_1), . \tag{1.5}$$

where

$$d_1 = \frac{\ln \frac{S_t}{K} + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}},$$

$$d_2 = d_1 - \sigma\sqrt{T_{\mathrm{T}} t}$$

and r still represents the risk free and $N(\cdot)$ is the Normal Cumulative Density Function.

As anticipated earlier mathematical tools are used to obtain approximations and in particular to be able to price derivatives of a financial underlying. However, the Black and Scholes model is inconsistent with empirical data mainly for two reasons regarding:

- absence of volatility smile curve
- asymmetry of log-return distributions

In order to talk about the absence of Smile curve of volatility, we must first introduce the concept of in implied volatility. When talking about European options the parameters of the model are the underlying current price, the strike, the maturity, the risk free interest rate and the volatility. Having fixed the reference asset, call prices in the market vary

[1] Itô's Lemma: used in stochastic calculus in order to compute the differential of a function of a particular type of stochastic process.

depending on the maturity and strike; it was observed that taking two different call contracts in the market (with different strikes and maturities observable in the market), and inverting the formula of the model to derive the volatility (implied, hence *implied volatility*), produces different values of σ, in contrast to the model's assumption that volatility, was a constant variable. To summarize, if we observe the implied volatility $\sigma(K,T)$, we do not find a constant value (as the theory would suggest) but a curve [37, 39], called volatility smile, which reflects the unsuitability of the model.

The second reason why the BS model is incorrect concerns the log-return distribution of the underlying asset. Considering the time series of the stocks prices, the log-return distribution of the daily price can be deduced:

$$\log\left(\frac{S_t}{S_0}\right) \sim \mathcal{N}\left(\mu t, \sigma^2 t\right) \quad \text{and} \quad \log\left(\frac{S_{t+\Delta}}{S_t}\right) \sim \mathcal{N}\left(\mu\Delta; \sigma^2\Delta\right)$$

where Δ indicates the time increment equal to one day (daily log-return). It was observed that the daily distribution of prices does not appear to be stable, so the Gaussian cannot represent this pattern: a correct model would have thicker tails since the Gaussian underestimate large movements.

Given the inability of the BS model to represent certain empirical dynamics, quantitative finance moved mainly in three directions:

- Local volatility model: Dupire in 1994 [15] introduced a model in which σ was not constant, but was seen as a deterministic function of the time and the price of the underlying asset:
$$dS_t = rS_t dt + \sigma\left(t, S_t\right) S_t dW_t \tag{1.6}$$

- Stochastic volatility model: the first model in this category was introduced by Hull and White in 1987 [26, 27], but more generally models in which volatility has its own stochastic markovian process belong to this class:
$$S_t = rS_t dt + \sigma_t S_t dW_t \tag{1.7}$$
where
$$\sigma_t = f(y_t) \qquad dy_t = \alpha\left(t, y_t\right) dt + \beta\left(t, y_t\right) d\tilde{W}_t \tag{1.8}$$

- Jumps model: consists of maintaining the dynamics based on a Markovian process, but adding the possibility of jumps in the price of the underlying asset by taking

advantage of Levy processes (X_t in the model):

$$S_t = S_0 e^{\mu t + X_t} \tag{1.9}$$

The process has no continuous trajectories; Merton in 1976 proposed the first diffusive jump model [35].

All of the above models were able to explain volatility smile curves in equity derivatives.

1.2. Local volatility model (LV)

Local volatility models assume that future volatility is a deterministic function of the underlying asset's value and time horizon (maturity); these characteristics are implicit in the current price of the vanilla option. Thus, the logic of the Black-Scholes model proved correct, and once the model was calibrated to the most liquid vanilla options in the market, it was possible to price and hedge exotic options. The model represented in Eq. (A.7), was introduced more generically by also considering the dividend rate (q):

$$S_t = (r - q)S_t dt + \sigma_t S_t dW_t \tag{1.10}$$

The key step in the local volatility framework is the Dupire formula [14, 15], through which local volatility can be linked to the implied volatility surface.

Theorem 1.1. ***Dupire's formula***
Let $C = C(K,T)$ be the price of a call option as a function of the strike K and of the time-to-maturity T. Then the local volatility function satisfies

$$\sigma_l^2(T,K) = \frac{\frac{\partial C}{\partial T} + (r-q)K\frac{\partial C}{\partial K} + qC}{\frac{K^2}{2}\frac{\partial^2 C}{\partial K^2}} \tag{1.11}$$

where r is the risk-free interest rate and q is the dividend rate.

From Eq. (1.11) it can be seen that for the calculation of local volatilities, it is necessary to consider the option's prime and second derivatives with respect to the strike and the prime derivatives with respect to the time-to-maturity: the numerical calculation of the derivatives may be unstable; consequently, it is essential to use a sufficiently smooth Black-Scholes implied volatility surface for the calculation of local volatility.

Even if the model is able to reproduce the implied volatility surface observed in the market, some weaknesses should be noted. Greeks calculated with a LV model are generally not

consistent with those observed empirically. It can also be seen that LV models tend to underestimate the volatility of volatility, evidence confirmed by the fact that structured products are often significantly underpriced. More in general, locare volatility models generate skew forwards too flat, leading to an inability to price certain products. Finally, a further limitation of the model is that the local volatility surface is composed by a relatively small number of points: therefore, to use unknown values, an approximation by interpolation must be made.

1.3. Stochastic volatility model (SV)

Stochastic volatility models make volatility aleatory by using a second markovian stochastic process to describe it. There are two Brownian motions in the model $\left(W_t, \tilde{W}_t\right)$, respectively, one referring to the equation for the price and the other for the volatility, whose correlation is represented by the parameter ρ. Eq. (1.7) referred to price seems to remain unchanged from the LV models,

$$S_t = rS_t dt + \sigma_t S_t dW_t$$

but the main difference occurs with the introduction of the dynamics process (Eq. (1.8))

$$\sigma_t = f(y_t) \qquad dy_t = \alpha\left(t, y_t\right) dt + \beta\left(t, y_t\right) d\tilde{W}_t$$

There are mainly three possibilities for the definition of this process:

- Geometric Brownian motion: Hull and White introduced the first model in this category [27], considering the dynamics of y_t as follows:

$$dy_t = c_1 y_t dt + c_2 y_t d\tilde{W}_t, \tag{1.12}$$

  where c_1 e c_2 are constants. The relation between y_t and σ is simply $\sigma = \sqrt{y_t}$ and the correlation between the two BM is zero (ρ=0). This model does not offer great results, in fact is the less used

- Ornstein-Uhlenbeck process [20]: the model introduced by Scott and Chesney [8, 38] is crucially important as one of the first models to introduce the concept of mean reverting, that is the tendency of a stochastic process to move back to its mean. They are introduced the variable λ, index of the speed of mean reversion, and η, the mean itself:

$$dy_t = \lambda\left(\eta - y_t\right) dt + \beta d\tilde{W}_t \tag{1.13}$$

Defining the drift as above, we see that the process tends to decrease when positive ($y_t > \eta$), and to increase when negative ($y_t < \eta$). Since the probability of y_t being negative is nonzero, we define $\sigma = e^{y_t}$; $\rho = 0$ as before.

- Cox-Ingersoll-Roll (CIR) process [11]: thanks to this process of defines the Heston model, the most important and used of this class [26]:

$$dy_t = \lambda(\eta - y_t)\,dt + \beta\sqrt{y_t}d\tilde{W}_t \tag{1.14}$$

which is a mean reverting and non-negative if y_0 and η are positive. In this case the correlation ρ is negative: this phenomenon is called *"Leverage Effect"* and indicates that when price decreases volatility tends to increase.

Heston model, in part because it can reproduce the Leverage Effect, is the most commonly used among the stochastic volatility models and has actually become a benchmark in derivative pricing and risk management.

1.4. Jumps model

The third model that researchers have proposed as an alternative to BMS to mainly explain the phenomenon of volatility smile/skew is the Jump model as

$$S_t = S_0 e^{\mu t + X_t}, \tag{1.15}$$

where X_t represents the Levy processes, fundamental in this theory. In fact, the Levy process is defined as the log-price of the underlying ($X_t \sim \log(S_t)$) and it is therefore crucial that we require that:

- X_t never goes to infinity, hence is limited, like S_t.
- X_t has discontinuous trajectories, since we are introducing jumps.

The Levi process family turns out to be large and complex, but by focusing on financial application this class results to be simpler. The first subclasses that can be highlighted are that of "Continuous Processes", of which the classic Black & Scholes model is part, and the one of "Discontinuous Processes", a real new development introduced. When talking about tje second class a further division is given:

- Finite Activity: In this category the jumps are a limited number and of varying

sizes. The most analyzed are the "Jump diffusion process" rappresented as:

$$X(t) = bt + \sigma W_t + \sum_{i=1}^{N(t)} Y_i, \tag{1.16}$$

where the Poisson process N_t represents when jumps occur and Y_i are i.i.d. random variables which can have different laws; These Levy processes are adequate to describe the log-prices, and according to the law of the random variables of the family (Y_i) there are two possibilities, which are the Merton [35] and the Kou model [32].

- Infinite Activity: they are characterised by an infinite number of infinitesimal jumps; the most commonly used models in this category are the Variance Gamma (VG) model [33] (and Extended VG, obtain by adding a Brownian Motion to the previous model) and the Normal-inverse Gaussian (NIG) model [3] (and Extended NIG, obtain by adding a Brownian Motion to the previous model).

All the models that have been listed in the previous sections were formulated with the mainly intention of trying to satisfy the implied volatility surface imposed by the market. It is well known that models differ in the very definition of volatility, but once calibrated almost all of them led to the same result. It was empirically observed that they recreated a volatility surface that was accurate in the short run, but in the long run inconsistent with market values. To solve this problem, new conditions and principles were introduced. This aims to create more innovative and satisfactory models.

1.5. Purposes for a new model

In order to move forward in the development of increasingly sophisticated financial models consistent with empirical data observable from the market, three characteristics to be satisfied, mainly related to volatility, are introduced. These features, already mentioned in the previous chapter, create a guideline for a new model introduced in the next section.

- Leverage Effect [6]: in the context of the stock world it refers to the tendency for stocks (and other financial instruments) to show a negative correlation between returns (returns) and volatility. This means that when stock market volatility increases, returns tend to decrease and vice versa. It means that when stock markets are under pressure and volatility is high, investors tend to see a reduction in returns, while when volatility is low, returns tend to be higher. The Leverage Effect can be particularly pronounced during volatile market periods or difficult economic conditions, when fear and uncertainty can lead to large swings in stock prices. In

summary, the Leverage Effect, as observed of time series, represents the negative correlation between returns and stock volatility, which can influence investment.

- Zumbach Effect [42]: it refers to the phenomenon of the dependence of volatility on past returns in financial markets. It is often observed in the context of time series analysis of financial data, particularly when examining the relationship between past returns and subsequent volatility. The Zumbach Effect is essentially a way of describing the correlation or relationship between past returns and future volatility; it implies that the level of volatility in a financial time series can be influenced by the recent history of returns: when a financial instrument experiences periods of high volatility, it is more likely to continue to have high volatility in the near future. Similarly, if a financial instrument has been relatively stable in terms of price movement, it is more likely to continue to be stable.

- Joint calibration of SPX and VIX implied volatility smiles: it is a process in financial modeling where both the options on the SPX (S&P 500 Index: equity index made up of 500 of the largest publicly traded companies in the United States selected by Standard & Poor's company) and options on the VIX (CBOE Volatility Index: financial benchmark that measures the market's expectations for future volatility) are considered simultaneously to determine the parameters of a mathematical model that can accurately represent the implied volatility curves of these two indices. The SPX implied volatility smile represents how implied volatility varies across different strike prices and maturities for options on the S&P 500. The shape of this smile can provide insights into market expectations for future market volatility and can be an essential input for pricing and risk management. Instead, the VIX, often referred to as the "fear gauge," is a measure of implied volatility for options on the S&P 500. The VIX implied volatility smile represents the implied volatility levels for different expiration dates on VIX options. Joint calibration refers to the process of fitting a mathematical model to the observed implied volatility smiles for both SPX and VIX options simultaneously. The goal is to find a model (such as a stochastic volatility model) that can accurately represent the dynamics of implied volatility for both indices and the complex dynamics of volatility in the market. This process involve estimating various model parameters that govern how implied volatility evolves over time and how it varies with different strike prices and maturities. Accurate modeling of implied volatility, may involving the use of historical option price data, is crucial for options pricing, risk management, and the development of investment strategies that involve SPX and VIX options. For a detailed description of VIX and joint calibration, see, for example [16] and [17] by Chicago Board Options Exchange or

[40] by Yingzi Zhu et al.

The Path-Dependent Volatility model ripsonde to these requirements.

2 | Path-Dependent Volatility models

Stochastic volatility (SV) models offered a framework for considering volatility as a stochastic process influenced by its own sources of randomness, which can be related to those governing asset price dynamics. After studies revealed that financial markets are characterized by a volatility dependence of the path itself of the underlying asset, a model was introduced by Guyon and Lekeufack to replicate this phenomenon: the "Path-Dependent Volatility" models (PDV) [25]. The objective of this chapter is to clarify how volatility is intrinsically linked to the path followed by asset returns. The development of the different models that led to the requirements for Leverage Effect, the Zumbach Effect and the Joint calibration of SPX and VIX implied volatility smiles will be observed, starting from the simple idea that:

$$\frac{dS_t}{S_t} = rdt + \sigma\left((S_u)_{u\leq t}\right) dW_t, \tag{2.1}$$

where S_t is the asset price, W_t is a standard Brownian motion and σ is the volatility.

2.1. Reasons for the PDV model

2.1.1. A market and modeling reason

In the area of volatility modeling, we run into two fundamental parameters with inherent scales: volatility levels and asset returns. Therefore, a valid model should establish a connection between these two aspects: PDV models fulfill this objective by clarifying the current level of volatility through historical asset returns.

We begin by considering that asset prices are not attainable within a predefined range; in fact, there are stocks with very different nominal values in the market. Different thinking regarding the scale of asset re turns and the scale of asset volatility. Indeed, it can be said that asset returns typically fluctuate by a few percentage points per day, except for

anomalies. Similarly, volatility remains stable regardless of the asset price. Considering the VIX index, it can be seen that over the years the value has remained relatively constant within the range [10, 80], regardless of the value of the SPX (Figure 2.1). It has been studied that the large variations the worked index, always maintained within the range, is due to significant fluctuations in the performance of the value of the SPX: for the same value of the SPX, different values of the VIX can be observed.

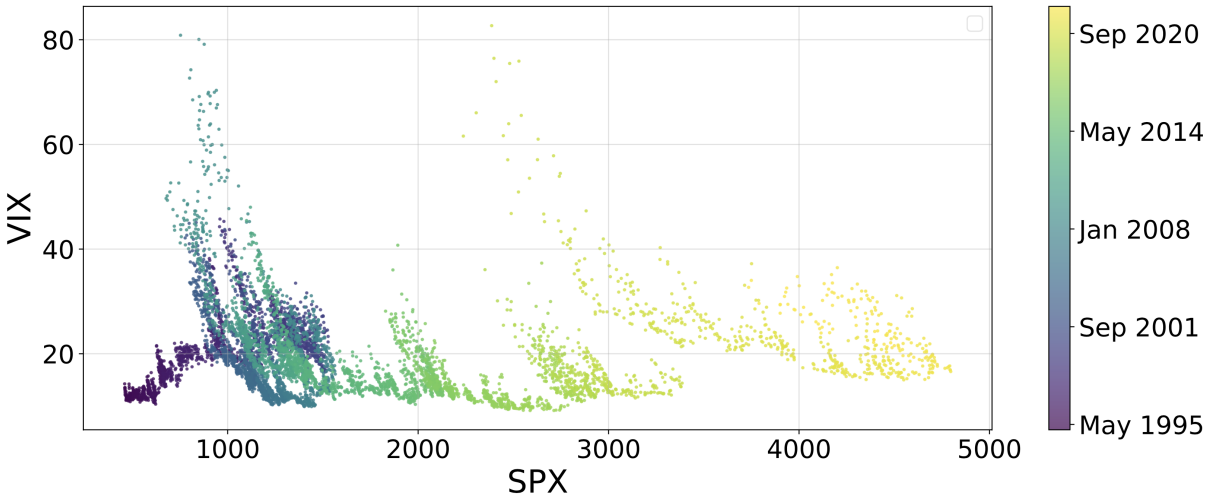


Figure 2.1: VIX vs SPX, January 1995–May 2022

This shows promptly that volatility is more accurately described by the historical behavior of the asset's price rather than its current valuation. Figure 2.1 provides further confirmation of the inadequacy of the relationship between volatility and asset price (typical of the LV model).

By analyzing temporal evolution in the historical data (Figure 2.2), it is clear that when the SPX follows an uptrend, the VIX shows a generally downward trend relative to the SPX, in line with the LV model. However, it is important to note that in the LV model, there is a kind of constraint on the reference level of the SPX when the SPX undergoes a sudden descent.

In general,we can take a closer look at the strengths of the PDV model compared with the LV and SV approaches. Let us begin with the Local Volatility model: the LV model attempts to explain the level of volatility as a function of asset level, which, as we have observed from historical data, does not have a solid financial background. Since PDV models accurately capture the observed dynamics of market volatility, they do not require recalibration as frequently as the LV model. For Stochastic Volatility models, there is evidence of a relationship between the variance of volatility and asset return, driven by

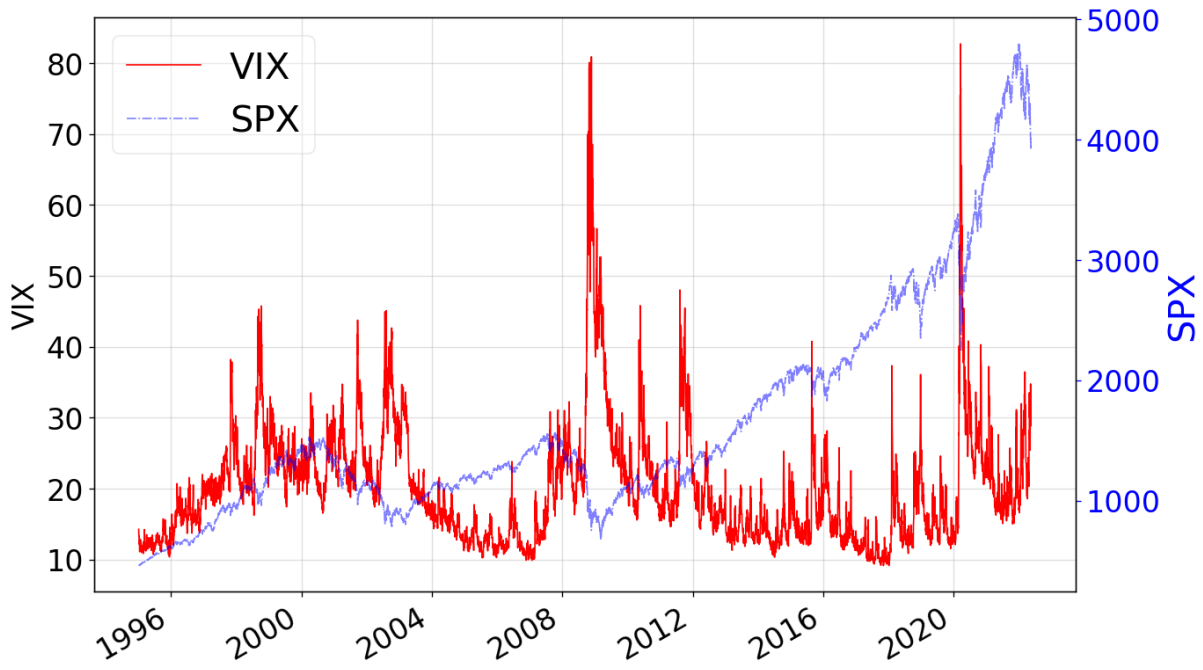


Figure 2.2: SPX and VIX time series, Jan 1, 1995–May 15, 2022

Brownian motions that influence the dynamics of asset prices and volatility. Although this approach makes more sense than LV because it uses asset returns as explanatory variables rather than asset levels, it may not be the most cost-effective choice. This is because it introduces path-dependence that is complex, is not explicitly selected, and may not reflect the price feedback effect on volatility as accurately as an explicitly chosen PDV model.

2.1.2. Role of Path-Dependent Volatility in Derivatives Pricing

In contrast to Stochastic Volatility (SV) models, Path-Dependent Volatility models (PDV) do not necessitate the introduction of supplementary sources of randomness, typically represented by Brownian motions, to generate intricate spot-volatility dynamics. PDV models elucidate volatility solely from an endogenous perspective, consequently, PDV models, unlike SV models, constitute comprehensive models where the pricing of derivatives is achieved in a unique and unambiguous manner, independent of any specific preference or utility function. SV models can attain completeness if volatility is permitted to rely on derivative prices of the underlying asset, however, establishing the joint dynamics of the asset's price and the prices of options linked to that asset is known to be exceedingly intricate. In PDV models, all information exchanged among market participants is encapsulated within the prices of the underlying asset, encompassing not just the current prices

but also the historical record of all past prices. While reality may exhibit some additional complexity, PDV models closely approximate it, so it is logical to initiate model construction by extracting the entirety of the volatility information contained in historical asset prices. In fact, PDV models are universally applicable to option pricing: all SV models possess an equivalent PDV counterpart, implying that the prices of all path-dependent options (not limited to vanilla options) on the underlying asset remain consistent across both models. This assertion is substantiated by Brunick and Shreve's demonstration in [7]: for a general Itô process given by $dS_t = \sigma_t S_t dW_t$, there exists a PDV model described as $d\hat{S}_t = \sigma(t, (\hat{S}_u)_{u \leq t})\hat{S}_t d\hat{W}_t$, where the distributions of the processes $(dS_t)_t \geq 0$ and $(\hat{S}_t)_t \geq 0$ are identical.

Finally, the objective need to support a path-dependency model with respect to volatility has already been presented in several academic papers, including a substantial part of the literature on Generalized Autoregressive Conditional Heteroskedasticity (GARCH) [5, 24], considered the most interesting and advance model till now thanks the representation of the Zumbach Effect. In SV models, a partial Zumbach Effect can be obtained [41, 42]; in particular, through a nonzero correlation between spot volatility and past volatilities realized on a larger scale, it is observed that past volatilities on a smaller scale are more highly correlated with future volatilities realized on a smaller scale, a result that PDV models also replicate, while going further. In fact, the innovative model introduced here fully deals with the Zumbach Effect by revealing that the dynamics of conditional volatility, relative to past observations, is influenced not only by the historical trajectory of past volatility but also by the historical path of prices, implying a dependence on the price path within the dynamics of volatility.

2.2. Volatility modeling

In this section we aim to understand the link between the level of volatility and past returns of the underlying asset and compare it with the previously developed relations.

2.2.1. Fundamental properties

Guyon and Lekeufack started the study of the new model from the path-dependent features that have the potential to explain most of the variability in volatility, focusing mainly on two of them, one related mainly to the recent trend of the underlying and the other to recent volatility activity:

- **Trend features**: The purpose of these features is to explain the fact that volatility,

in the presence of a decline in asset value, tends to increase (Leverage Effect). To do this, a weighted sum of past daily returns is introduced:

$$R_{1,t} := \sum_{t_i \leq t} K_1 (t - t_i) r_{t_i} \tag{2.2}$$

where

$$r_{t_i} := \frac{S_{t_i} - S_{t_{i-1}}}{S_{t_{i-1}}} \tag{2.3}$$

represents the daily return. K_1 represents a convolution kernel that assigns varying degrees of significance to historical daily returns, depending on how distant in the past they were observed. As time elapses, the kernel gradually diminishes in value, signifying the diminishing influence of past daily returns.

- **Activity features**: Activity characteristics are those that reflect recent fluctuations in the price of an asset, regardless of its general direction. These characteristics prove crucial in the process of modeling the phenomenon known as volatility clustering, where periods characterized by high volatility tend to be followed by other periods of high volatility, and periods of low volatility usually precede other periods of low volatility (Zumbach Effect). Another manifestation of volatility clustering is evidenced in the observation that implied volatility tends to increase when historical volatility is higher. An important example of a volatility-related feature is the weighted sum of past daily returns squared:

$$R_{2,t} := \sum_{t_i \leq t} K_2 (t - t_i) r_{t_i}^2 \tag{2.4}$$

where K_2, as before, is a convolution kernel that gradually decreases in value over time. In order to take into account even higher moments of past daily returns, the K_2-weighted historical volatility is used: $\sum_t := \sqrt{R_{2,t}}$.

2.2.2. The Volatility model

Guyon and Lekeufack's model, based on the considerations mentioned earlier, explains the instantaneous volatility, and not its square the variance, as a simple affine combination of the features:

$$\sigma_t = \beta_0 + \beta_1 R_{1,t} + \beta_2 \Sigma_t, \quad \beta_0 > 0, \quad \beta_1 < 0, \quad \beta_2 \in (0, 1) \tag{2.5}$$

The "trend feature" $R_{1,t}$ and the "historical volatility features" $\sum_t$ are related to each other by coefficients that, in order to guarantee particular properties to the volatility, are subject to limitations:

- $\beta_1 < 0$ produces Leverage Effect because there is a negative linear dependence on past trends in yields : negative trends will have a positive impact on the current level of volatility, while positive trends will have a negative effect.
- $\beta_2 > 0$ introduces a positive correlation with historical volatility, enabling the replication of volatility clustering and indirectly explains the Zumbach Effect.
- $\beta_1 < 1$ ensures the stability of the model, as in GARCH model.

Talking about limitations, it is necessary to specify what are the appropriate types of kernel (K_1, K_2) acceptable for the model. The two kernels, defined for different model features are distinct in one from the other, even if they have the same purpose of weighting short-to-long memory events. We want to model these two kernels in such a way that the weights $K_n(\tau)$ decrease fast for small time lags τ (recent events), but simultaneously decrease for large time lags (more dated events); in this way volatility can have short and long memory. Guyon and Lekeufack investigated which was, empirically, the best choice, selecting a kernel that resembles a power law, except that for null delays τ the kernels do not seem to explode, thus both kernels were chosen as time-shifted power laws (TSPL):

$$K(\tau) = K_{\alpha,\delta}(\tau) := Z_{\alpha,\delta}^{-1}(\tau + \delta)^{-\alpha}, \quad \tau \geq 0, \quad \alpha > 1, \quad \delta > 0 \tag{2.6}$$

The time shift δ guarantees that $K_{\alpha,\delta}(\tau)$ does not blow up when the lag τ vanishes, but we will select all positive values. $Z_{\alpha,\delta}$ is a normalization coefficient that allows the assumptions that the sum of the weights is equal to one ($\sum_{i=0}^{+\infty} K(i\Delta t) = 1$) in both discrete-time and continuous-time situations where it turns out to be

$$Z_{\alpha,\delta} = \int_0^\infty (\tau + \delta)^{-\alpha} d\tau = \frac{\delta^{1-\alpha}}{\alpha - 1} \tag{2.7}$$

The inclusion of a small (but positive) time shift δ emphasizes the importance of recent returns (short-term memory) already identified by Gatheral, Jaisson, and Rosenbaum [18]. In addition, the model thus formulated succeeds in adding harmoniously to the observed long memory explanation of volatility identified by Comte and Renault [9].

On a practical level, however, TSPL kernels do not produce a Markovian model[1], and

[1] A Markovian model, is a mathematical framework used to describe systems or processes where the future state of the system depends only on its current state and is independent of its past states.

consequently are complicated and slow to simulate; however, there is an alternative. Exponential Kernels ($K^{(\lambda)}(\tau) := \lambda e^{-\lambda\tau}, \quad \lambda > 0$), introduced in the discrete-time version of the threshold EWMA Heston model, turn out to be Markovian models and thus easy to simulate even through the Monte Carlo Method, one of the most widely used techniques in financial pricing. It is shown and empirically observable that a convex combination of two exponential kernels (2-EXP), which mixes a very short time scale with a longer time scale, offers another very natural way to reconcile long memory, short memory, and volatility-like behavior. In fact, in Figure 2.3 it can be seen that a 2-EXP and a TSPL can approximate each other very well over a wide range of maturities. Although it is more intuitive to use TSPL kernels, an exponential kernel for K1 and one for k2 (2-factors PDV model) and later a convex combination of two exponential kernels for both K_1 and K_2 (4-factors PDV model) will be used initially in the development and simulations of continuous-time PDV models.

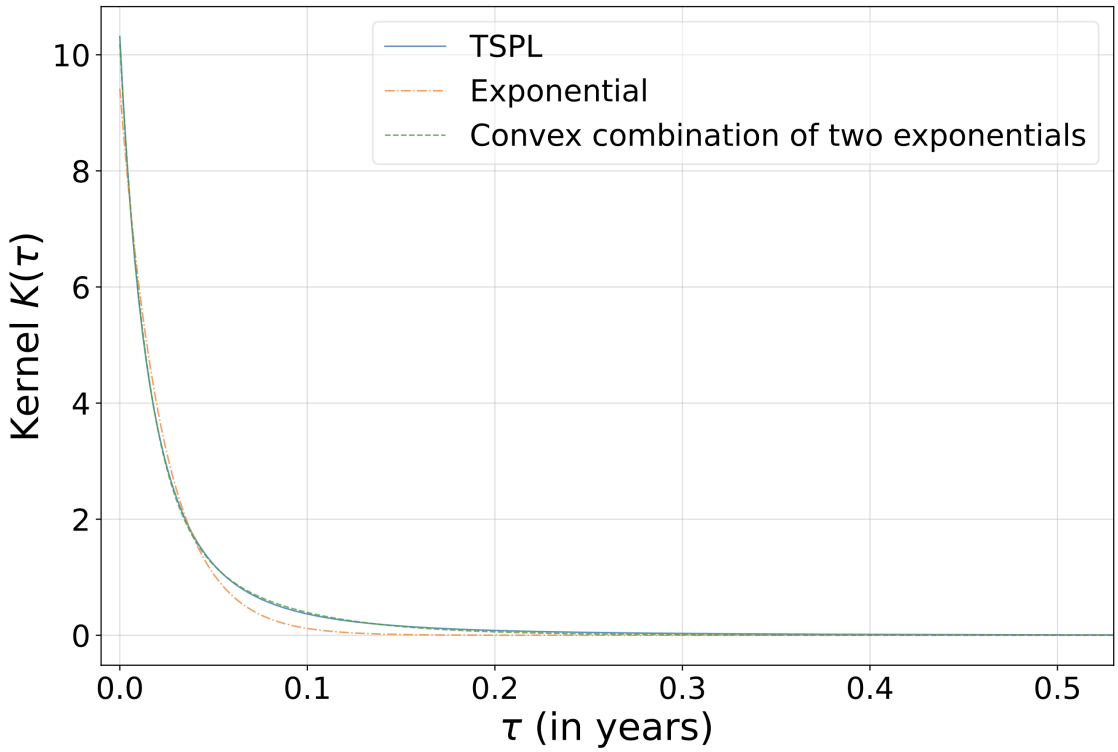


Figure 2.3: TSPL Kernel K_1 for the prediction of the Realized volatility (of SPX) and the best fits by one exponential and by a convex combination of two exponentials. The convex combination is a much better fit.

2.3. The model

We have established that through a simple linear model much of the observed variability in volatility is explained. Now we want to link the volatility model to one of a generic financial underlying asset S_t.

2.3.1. Continuous-Time Path-Dependent Volatility model

Thanks to Eq. (2.5) we can define a model that, by convention, neglects dividends and takes 0 as the interest rate:

$$\begin{aligned}
\frac{dS_t}{S_t} &= \sigma_t dW_t \\
\sigma_t &= \sigma\left(R_{1,t}, R_{2,t}\right) \\
\sigma\left(R_1, R_2\right) &= \beta_0 + \beta_1 R_1 + \beta_2\sqrt{R_2} \\
R_{1,t} &= \int_{-\infty}^{t} K_1(t-u)\frac{dS_u}{S_u} = \int_{-\infty}^{t} K_1(t-u)\sigma_u dW_u \\
R_{2,t} &= \int_{-\infty}^{t} K_2(t-u)\left(\frac{dS_u}{S_u}\right)^2 = \int_{-\infty}^{t} K_2(t-u)\sigma_u^2 du
\end{aligned} \tag{2.8}$$

where the dynamics of $R_{1,t}$ and $R_{1,t}$, which in general are not Markovian because of the general kernels (TSPL) are

$$\begin{aligned}
dR_{1,t} &= \left(\int_{-\infty}^{t} K_1'(t-u)\frac{dS_u}{S_u}\right)dt + K_1(0)\frac{dS_t}{S_t} = \left(\int_{-\infty}^{t} K_1'(t-u)\sigma_u dW_u\right)dt + K_1(0)\sigma_t dW_t \\
dR_{2,t} &= \left(\int_{-\infty}^{t} K_2'(t-u)\left(\frac{dS_u}{S_u}\right)^2\right)dt + K_2(0)\left(\frac{dS_t}{S_t}\right)^2 = \left(K_2(0)\sigma_t^2 + \int_{-\infty}^{t} K_2'(t-u)\sigma_u^2 du\right)dt
\end{aligned}$$

As mentioned earlier the model (2.8) is further developed by eliminating TSPL kernels and introducing exponential ones: this will allow additional expansion of the dynamics of $R_{1,t}$ and $R_{2,t}$.

2.3.2. 2-factor Markovian PDV model

The main changes from model (2.8) concern kernels, in fact they are chosen exponential type:

$$\begin{aligned}
K_1(\tau) &:= K^{(\lambda_1)}(\tau) := \lambda_1 e^{-\lambda_1\tau}, \quad \lambda_1 > 0 \\
K_2(\tau) &:= K^{(\lambda_2)}(\tau) := \lambda_2 e^{-\lambda_2\tau}, \quad \lambda_2 > 0
\end{aligned} \tag{2.9}$$

With the introduction of the exponential kernels, where l1 and l2 are normalization coefficients, the simplest Markovian model is obtained, and due to the very nature of the changes made, the integrals present in the dynamics of $R_{1,t}$ and $R_{2,t}$ can be developed. the model becomes:

$$\begin{aligned}
\frac{dS_t}{S_t} &= \sigma_t dW_t \\
\sigma_t &= \sigma\left(R_{1,t}, R_{2,t}\right) \\
\sigma\left(R_1, R_2\right) &= \beta_0 + \beta_1 R_1 + \beta_2 \sqrt{R_2} \\
dR_{1,t} &= \lambda_1 \left(\frac{dS_t}{S_t} - R_{1,t}dt\right) = \lambda_1 \left(\sigma\left(R_{1,t}, R_{2,t}\right) dW_t - R_{1,t}dt\right) \\
dR_{2,t} &= \lambda_2 \left(\left(\frac{dS_t}{S_t}\right)^2 - R_{2,t}dt\right) = \lambda_2 \left(\sigma\left(R_{1,t}, R_{2,t}\right)^2 - R_{2,t}\right) dt
\end{aligned} \tag{2.10}$$

Since the instantaneous volatility is a deterministic function of two markovian factors ($R_{1,t}$, $R_{2,t}$), this model has been called the 2-factor Markovian PDV model. As discussed in Section 2.2.2, the selection of K_1 and K_2 as single exponential kernels proves inadequate in capturing the blend of short-term and long-term memory present in both $R_{1,t}$ and $R_{2,t}$. While Model (2.10) may not offer a complete solution, its primary advantage lies in its simplicity. An analysis of this model yields valuable qualitative insights into the dynamics of the volatility startin fomr Eq. (2.5)[2] and applying Itô's Lemma[3]:

$$d\sigma_t = \left(-\beta_1\lambda_1 R_{1,t} + \frac{\beta_2\lambda_2}{2}\frac{\sigma_t^2 - R_{2,t}}{\sqrt{R_{2,t}}}\right) dt + \beta_1\lambda_1\sigma_t dW_t \tag{2.11}$$

The evolution of volatility introduces a certain level of complexity. Unlike the volatility of volatility, which is influenced solely by its own dynamics, the drift of volatility depends on the individual components, R1 and R2, unless $\beta_1 = 0$ or $\beta_2 = 0$. We will then examine the evolution of volatility in some particular situations in order to provide an explanation for the parameter ranges in the volatility model (2.5):

- $\beta_2 = 0$: the behavior of $\sigma_t = \beta_0 + \beta_1 R_{1,t}$ simplifies to the classical self-regulating diffusion process, described by the differential equation

$$d\sigma_t = \lambda_1\left(\beta_0 - \sigma_t\right) dt + \beta_1\lambda_1\sigma_t dW_t$$

  This implies that the instantaneous volatility σ_t tends to return to β_0 at a constant

[2]Reminder: $\sum_t := \sqrt{R_{2,t}}$
[3]Itô's Lemma: General application shown in the Appendix A.1.

rate λ_1, exhibiting a mean-reverting behavior.

- $\beta_1 = 0$: volatility dynamics becomes$\sigma_t = \beta_0 + \beta_2 R_{2,t}$, but its differential form still remains relatively complex:

$$d\sigma_t = \frac{\lambda_2 \left(1 - \beta_2^2\right)}{2} \frac{\sigma_t - \gamma}{\sigma_t - \beta_0} \left(\sigma^* - \sigma_t\right) dt,$$

where

$$\gamma := \frac{\beta_0}{1 + \beta_2}, \quad \sigma^* := \frac{\beta_0}{1 - \beta_2}$$

and in particular $\gamma < \beta_0 < \sigma^*$.

Therefore, in both cases, the volatility mean reverts. In the general case where $\beta_1 \neq 0$ and $\beta_2 \neq 0$, the drift μ_t of σ_t is not simply a function of the volatility itself but a function of the two factors $R_{1,t}$ and $R_{2,t}$,

$$\mu\left(R_1, R_2\right) := -\beta_1 \lambda_1 R_1 + \frac{\beta_2 \lambda_2}{2} \frac{\left(\beta_0 + \beta_1 R_1 + \beta_2 \sqrt{R_2}\right)^2 - R_2}{\sqrt{R_2}} \tag{2.12}$$

In the drift of the differential form of volatility, two primary components can be discerned: one influenced by $R_{1,t}$ and the other influenced by $R_{2,t}$. The first one, driven by its mean reversion to zero, leads to a rapid reduction in volatility. Conversely, the second component contributes to a gradual and sustained increase in volatility, albeit with less intensity. This phenomenon elucidates why, following a sharp spike in volatility that momentarily recedes, volatility can persist at relatively high levels for an extended period, even when the asset price undergoes a swift recovery. This persistence is attributed to the recollection of past volatility through squared returns, which competes with the influence of past signed returns.

Finally, in order not to have negative volatility, certain restrictions must be placed on the λ_1 and λ_2 parameters of the kernels. Considering the situation in which σ_t is zero, the diffusion term of the differential Eq. (3.48) vanishes ($\beta_1 \lambda_1 \sigma_t = 0$), but its drift term does not, and being the volatility a non-negative quantity, it must be imposed greater than

zero. So, considering the initial constraints on the parameters, we have

$$\begin{cases} \mu\left(R_1, R_2\right) = -\beta_1\lambda_1 R_{1,t} + \dfrac{\beta_2\lambda_2}{2}\dfrac{\sigma_t^2 - R_{2,t}}{\sqrt{R_{2,t}}} > 0 \\ \sigma_t = \beta_0 + \beta_1 R_{1,t} + \beta_2\sqrt{R_{2,t}} = 0 \\ R_{1,t}, R_{2,t} > 0 \\ \beta_0, \beta_2 > 0 \\ \beta_1 < 0 \\ \lambda_1, \lambda_2 > 0 \end{cases} \tag{2.13}$$

solving the system we obtain that $\lambda_2 < 2\lambda_1$: thus the drift of volatility is positive and consequently we have non-negative volatility

2.3.3. 4-factor Markovian PDV model

The 2-factor PDV model (2.10), although relatively simple, uses single exponential kernels for k1 and k2, which consequently fail to replicate the TSPL kernel. Here we introduce a convex combination of two exponential kernels for K_1 and K_2 since they better approximate TSPL kernels (see Figure 2.3):

$$\begin{aligned} K_1(\tau) &= (1-\theta_1)\,\lambda_{1,0}e^{-\lambda_{1,0}\tau} + \theta_1\lambda_{1,1}e^{-\lambda_{1,1}\tau} \\ K_2(\tau) &= (1-\theta_2)\,\lambda_{2,0}e^{-\lambda_{2,0}\tau} + \theta_2\lambda_{2,1}e^{-\lambda_{2,1}\tau} \end{aligned} \tag{2.14}$$

where $\theta_i \in [0,1], \lambda_{i,0} > \lambda_{i,1} > 0$ for $i \in \{1,2\}$. Thanks to the mixing factor θ we are able to give volatility short and long memory (Section 2.2.2) and to introduce a new 4-factor model:

$$\begin{aligned} \frac{dS_t}{S_t} &= \sigma_t dW_t \\ \sigma_t &= \sigma\left(R_{1,t}, R_{2,t}\right) \\ \sigma\left(R_1, R_2\right) &= \beta_0 + \beta_1 R_1 + \beta_2\sqrt{R_2} \\ R_{1,t} &= (1-\theta_1)\,R_{1,0,t} + \theta_1 R_{1,1,t} \\ R_{2,t} &= (1-\theta_2)\,R_{2,0,t} + \theta_2 R_{2,1,t} \end{aligned} \tag{2.15}$$

$$dR_{1,j,t} = \lambda_{1,j}\left(\frac{dS_t}{S_t} - R_{1,j,t}dt\right) = \lambda_{1,j}\left(\sigma\left(R_{1,t}, R_{2,t}\right)dW_t - R_{1,j,t}dt\right), \quad j \in \{0,1\},$$

$$dR_{2,j,t} = \lambda_{2,j}\left(\left(\frac{dS_t}{S_t}\right)^2 - R_{2,j,t}dt\right) = \lambda_{2,j}\left(\sigma\left(R_{1,t}, R_{2,t}\right)^2 - R_{2,j,t}\right)dt, \quad j \in \{0,1\}.$$

The model's additional parameters, θ_1 and θ_2, blend the influence of recent and historical returns (or squared returns) to create the summary random variables R_1 and R_2, respectively. It is possible to retrace the same reasoning used for the 2-factor PDV model in Section 2.3.2 with a few adjustments. We have to introduce, for $n \in \{1, 2\}$, the average quantities

$$\bar{\lambda}_n := (1 - \theta_n)\, \lambda_{n,0} + \theta_n \lambda_{n,1}$$
$$\bar{R}_{n,t} := \frac{(1 - \theta_n)\, \lambda_{n,0} R_{n,0,t} + \theta_n \lambda_{n,1} R_{n,1,t}}{\bar{\lambda}_n}$$

In this way, the dynamics of volatility can be deduced, again using Itô's Lemma:

$$d\sigma_t = \left(-\beta_1 \bar{\lambda}_1 \bar{R}_{1,t} + \frac{\beta_2 \bar{\lambda}_2}{2} \frac{\sigma_t^2 - \bar{R}_{2,t}}{\sqrt{R_{2,t}}} \right) dt + \beta_1 \bar{\lambda}_1 \sigma_t dW_t \tag{2.16}$$

The dynamics of this process exhibit certain qualitative properties that resemble the behavior of the asset price. Specifically, the drift of σ_t leads to the phenomenon of volatility clustering, owing to a discernible trend of mean reversion in volatility. Additionally, the lognormality of σ_t remains constant over time, as observed in the data. Furthermore, it's important to note that the dynamics of σ_t are influenced by the price path, making them price path-dependent. In other words, the drift of σ_t isn't solely determined by past values of volatility (σ_u) for $u \leq t$; it is also influenced by past asset returns, as reflected through $\bar{R}_{1,t}$.

Another remarkable property of the 4-factor PDV model is its ability to generate highly realistic smiles for both the SPX and the VIX (will be shown in the next chapters). In fact, it even allows for the joint calibration of the 4-factor PDV model to accurately match market SPX and VIX smiles. This demonstrates, for the first time, that a continuous-time Markovian parametric stochastic (in fact, path-dependent) volatility model can effectively address the joint SPX/VIX smile calibration challenge.

2.4. Path-Dependent Stochastic Volatility

Volatility is primarily driven by historical paths, yet it's essential to note that it's not exclusively determined by past movements, idea on which the previous models are based. Anomalous returns can occur, even when considering path-dependent volatility, due to unforeseen news or newly acquired information. To address these unusual returns, which represent the exogenous component of volatility, we propose the Path-Dependent Stochastic Volatility (PDSV) model. This model resembles the 4-factor PDV model, with one crucial difference: the instantaneous volatility $\sigma_t = a_t \sigma\left(R_{1,t}, R_{2,t}\right)$ is now the result of

the product of PDV $\sigma(R_{1,t}, R_{2,t})$ and a stochastic volatility component (a_t). The model becomes:

$$\begin{aligned} \frac{dS_t}{S_t} &= \sigma_t dW_t \\ \sigma_t &= a_t \sigma(R_{1,t}, R_{2,t}) \\ \sigma(R_1, R_2) &= \beta_0 + \beta_1 R_1 + \beta_2 \sqrt{R_2} \\ R_{1,t} &= (1-\theta_1) R_{1,0,t} + \theta_1 R_{1,1,t} \\ R_{2,t} &= (1-\theta_2) R_{2,0,t} + \theta_2 R_{2,1,t} \end{aligned} \tag{2.17}$$

$$dR_{1,j,t} = \lambda_{1,j}\left(\frac{dS_t}{S_t} - R_{1,j,t}dt\right) = \lambda_{1,j}\left(a_t\sigma(R_{1,t}, R_{2,t})\, dW_t - R_{1,j,t}dt\right), \quad j \in \{0,1\},$$

$$dR_{2,j,t} = \lambda_{2,j}\left(\left(\frac{dS_t}{S_t}\right)^2 - R_{2,j,t}dt\right) = \lambda_{2,j}\left(a_t^2\sigma(R_{1,t}, R_{2,t})^2 - R_{2,j,t}\right)dt, \quad j \in \{0,1\}.$$

where the stochastic process added (a_t) is an Ornstein-Uhlenbeck (OU) process

$$da_t = \kappa(1 - a_t)\, dt + \nu dZ_t, \tag{2.18}$$

with $\nu, \kappa > 0$ and Z a Brownian Motion or an exponential-OU process

$$a_t = e^{X_t}, \quad dX_t = -\kappa X_t dt + \nu dZ_t. \tag{2.19}$$

Because of the supplementary random process added the model is no more pure-feedback; moreover we can consider it a 5-factor Markovian model (easy and fast to simulate).

3 | Practical applications

In this Chapter we will go on to deal with the methodologies that bring the theoretical model described in the previous chapter to practical applications in derivatives computation. In particular, we will study two pricing techniques. The first exploits the Monte Carlo method, which is well-known in the financial branch and simple to apply. The second, however, is via binomial tree, which is more complex to model and the actual innovation of this Thesis. Both techniques require discretization of the model, which can be done in several ways described later. Thus, the purpose of this chapter is to explain how to go from a theoretical model to a model that can be used in reality thanks to computational platforms.

3.1. Monte Carlo method

In the financial sector, calculating the price of derivatives analytically and accurately is often difficult due to the many sources of randomness in the models. The Monte Carlo method is a suitable solution to this problem [36]. In fact, the Monte Carlo method (MC) is based on the principle of simulations: the same calculation is replicated several times by assigning different random values to the aleatory variables chosen according to the distribution describing them in the model. This method therefore does not turn out to be exact, but we will show that results obtained have very small confidence intervals. As anticipated in the previous chapter, the choice of kernels in the PDV model leads to a markovian model, a model in which the future state of the system depends only on its current state and is independent of its past states, and thus the Monte Carlo method turns out to be applicable.

3.1.1. Theory for application to vanilla options

Monte Carlo simulations are convenient since prices can be written as expected values, which can indeed be computed exploiting simulations [21]. We start from the theory of partial derivative equations (PDE [31]) related to vanilla options, to then develop the one with respect to Monte Carlo, via the Feyman Kac formula [12], all in the context

of a risk-neutral probability measure (Q-measure).[1] We consider a generic underlying S_t modelled as follows:

$$S_t = S_0 e^{rt+X(T)} \tag{3.1}$$

where $X(t)$ is an Itô process driven under the risk-free measure by the following dynamics

$$dX_t = a\left(X_t, t\right) dt + b\left(X_t, t\right) dW_t \quad X_0 = x \tag{3.2}$$

(x is directly related to the value of the underlying $S_0 = s$). The purpose is to understand how for a generic contingent claim, the price of the derivative can be obtained by knowing the maturity T, the value of the underline, the payoff function ϕ (call, put, etc.) and the interest rate ($k(t)$ seen as a deterministic interest, then set equal to zero for simplicity). The solving formula is

$$v(x,t) = \mathbb{E}^Q \left[e^{-\int_t^T k(u)du} \phi\left(X_T\right) \mid X_t = x \right] \tag{3.3}$$

if we look at the log-price X_t or

$$F(s,t) = \mathbb{E}^Q \left[e^{-\int_t^T k(u)du} \phi\left(S_T\right) \mid S_t = s \right] \tag{3.4}$$

if we look at the price S_t; it is the result of solving the PDE, analyzed by the Feyman Kac formula.

To prove the formula about the log-price X_t, we consider the following PDE and the terminal condition:

$$\begin{cases} \partial_t v(x,t) + a(x,t)\partial_x v(x,t) + \frac{1}{2} b^2(x,t)\partial^2_{xx} v(x,t) - k(t)v(x,t) = 0 \\ v(x,T) = \phi(x) \end{cases} \tag{3.5}$$

For the sake of simplicity we set the risk free $k(t) = 0$. Applying the Itô's Lemma to $v(x(t), t)$ where $dx(t) = a(x,t)dt + b(x,t)dW_t$ we get:

$$dv = \left(\partial_t v + a\partial_x v + \frac{1}{2} b^2 \partial^2_{xx} v \right) dt + b\partial_x v dW_t \tag{3.6}$$

where the drift part ($\partial_t v + a\partial_x v + \frac{1}{2} b^2 \partial^2_{xx} v$) is equal to 0 according to the PDE (3.5).

[1] Particular probability measure that guarantees the absence of arbitrage possibilities in the model considered.

Let us now integrate:

$$\int_t^T dv = v\left(X_T, T\right) - v(x,t) = \int_t^T b(x,t)\partial_x v(x,t) dW_t$$
$$\Rightarrow v(x,t) = v\left(X_T, T\right) - \int_t^T b(x,t)\partial_x v(x,t) dW_t$$

and computing the expectation on both sides:

$$\mathbb{E}[v(x,t)] = \mathbb{E}\left[v\left(X_T, T\right) - \int_t^T b(x,t)\partial_x v(x,t) dW_t\right]$$
$$v(x,t) = \mathbb{E}\left[v\left(X_T, T\right)\right]$$

since that the expectation of an Itô integral with respect to a Wiener process is always null, so:

$$v(x,t) = \mathbb{E}\left[v\left(X_T, T\right)\right] = \mathbb{E}\left[\phi\left(X_T, T\right)\right] = \mathbb{E}\left[\phi\left(X_T, T\right) \mid X_t = x\right] \tag{3.7}$$

In order to apply this formula to the models presented in the previous chapter (2-factor, 4-factor, 5-factor) some considerations must be made. It is necessary to keep in mind that when dealing with the models described, the drift and diffusion term of the stochastic differential equation (3.2) are referred to as vectors, since in the 2-factor model the dynamics to be described are 3 (X_t, R_1 and R_2) and in the 4-factor model they are 5 (X_t, $R_{1,0}$, $R_{1,1}$, $R_{2,0}$ and $R_{2,1}$). The parameters in these vectors must be selected in such a way as to ensure the existence, uniqueness and boundedness of the solution of the SDE.

3.1.2. Considerations on Monte Carlo simulations

Now we want to understand the accuracy of the results obtained by the Monte Carlo method, since so far we have focused exclusively on analytical reflections. Now, let's explore the numerical estimation of the expected value of a specific random variable, X, using the Monte Carlo method [28]. We can represent this expected value with the symbol θ, and its estimator is as follows:

$$\hat{\theta}_n = \frac{\sum_{i=1}^n X^{(i)}}{n} \tag{3.8}$$

where $X = \left(X^{(1)}, \ldots, X^{(n)}\right)$ is a random vector with i.i.d. (same distribution for all the variables). First, it is important to emphasize that $\hat{\theta}_n$ is an estimate of θ. As $\hat{\theta}_n$ has an expected value of θ and a variance of Var(X)/n, as n increases, $\hat{\theta}_n$ gradually approaches the true value of θ; this is affirmed by the law of large numbers.

Theorem 3.1. ***Law of large numbers***
Let $(X_i)_{i>0}$ be a set of i.i.d. random variable with mean μ. If we define:

$$Y(n) = \frac{1}{n}\sum_{i=1}^{n} X_i$$

than

$$\lim_{n\to+\infty} P(|Y(n) - \mu| \geq \eta) = 0 \quad \forall \eta > 0$$

Second, we want to evaluate the quality of the approximation provided by $\hat{\theta}_n$. For this purpose, it is necessary to introduce another theorem: the Central Limit Theorem

Theorem 3.2. ***Central Limit Theorem***
Let $(X_i)_{i>0}$ be a set of i.i.d. random variables with mean μ and variance σ^2. Then:

$$\lim_{n\to+\infty} \frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \sim \mathcal{N}(0,1)$$

As a consequence:

$$\lim_{n\to+\infty} \frac{\hat{\theta}_n - n\theta}{\sigma\sqrt{n}} \sim \mathcal{N}(0,1)$$
$$\Rightarrow \lim_{n\to+\infty} \left(\hat{\theta}_n - \theta\right) \sim \mathcal{N}\left(0, \sigma^2/n\right)$$

The Monte Carlo error follows a normal distribution with a mean of zero and a standard deviation of $\frac{\sigma}{\sqrt{n}}$, which is often referred to as the standard error of MC. Consequently, if we aim to decrease the error by a factor of N, we must increase the number of simulations, denoted as n, by a factor of N^2. This demonstrates that Monte Carlo simulations can be computationally intensive.

The situation becomes more complicated considering that the variance of the random variable, σ, is unknown and consequently it is necessary to use an estimate for it to compute the approximation of the standard error, which becomes $\frac{\hat{\sigma}}{\sqrt{n}}$. The normalized estimator, $\left(\hat{\theta}_n - \theta\right)/(\hat{\sigma}/\sqrt{n})$, converges in probability to a standard normal distribution. As a result, the cumulative distribution function of this normalized estimator converges to the standard cumulative normal distribution, denoted as $\Phi(x)$:

$$P\left(\frac{\hat{\theta}_n - \theta}{\hat{\sigma}/\sqrt{n}} < x\right) \underset{n\to+\infty}{\longrightarrow} \Phi(x)$$

Now it is possible to estimate the probability that θ is in the interval $\left(\hat{\theta}_n - a\frac{\hat{\sigma}}{\sqrt{n}}; \hat{\theta}_n + b\frac{\hat{\sigma}}{\sqrt{n}}\right)$:

$$P\left(\theta \in \left(\hat{\theta}_n - a\frac{\hat{\sigma}}{\sqrt{n}}; \hat{\theta}_n + b\frac{\hat{\sigma}}{\sqrt{n}}\right)\right)$$
$$P\left(\theta < \hat{\theta}_n + b\frac{\hat{\sigma}}{\sqrt{n}}\right) - P\left(\theta \leq \hat{\theta}_n - a\frac{\hat{\sigma}}{\sqrt{n}}\right)$$

Considering that

$$P\left(\frac{\hat{\theta}_n - \theta}{\hat{\sigma}/\sqrt{n}} < -a\right) = \Phi(-a)$$

and that

$$P\left(\frac{\hat{\theta}_n - \theta}{\hat{\sigma}/\sqrt{n}} < b\right) = P\left(\frac{\hat{\theta}_n - \theta}{\hat{\sigma}/\sqrt{n}} > -b\right) = 1 - P\left(\frac{\hat{\theta}_n - \theta}{\hat{\sigma}/\sqrt{n}} \leq -b\right) = 1 - \Phi(-b) = \Phi(b)$$

we have

$$P\left(\theta \in \left(\hat{\theta}_n - a\frac{\hat{\sigma}}{\sqrt{n}}; \hat{\theta}_n + b\frac{\hat{\sigma}}{\sqrt{n}}\right)\right) \underset{n\to+\infty}{\longrightarrow} \Phi(b) - \Phi(-a)$$

Moreover, in order to have a symmetric confidence interval we take $a = b = z_{\delta/2}$ where $\Phi\left(z_{\delta/2}\right) = 1 - \delta/2$: in this case indeed $\Phi\left(z_{\delta/2}\right) - \Phi\left(-z_{\delta/2}\right) = \Phi\left(z_{\delta/2}\right) - \left(1 - \Phi\left(z_{\delta/2}\right)\right) = 1 - \delta$ and δ is set according to the confidence we want to have; in the results shown in the next chapter, we will consider a 95% confidence interval.

3.1.3. Variance reduction techniques

In the previous section, we delved into the application of Monte Carlo simulations for estimating θ, that represents an estimate for the price of the vanilla option, and its corresponding confidence interval, which can be represented as:

$$\left(\hat{\theta}_n - z_{\delta/2}\frac{\hat{\sigma}}{\sqrt{n}}; \hat{\theta}_n + z_{\delta/2}\frac{\hat{\sigma}}{\sqrt{n}}\right) \tag{3.9}$$

Here, the value of $z_{\delta/2}$ is determined by the desired level of confidence and is independent of the specific problem. To enhance the precision of our estimate and narrow the confidence interval, we can either increase the number of simulations, or apply variance reduction techniques to $\hat{\sigma}$ [22].

Let's consider two distinct random variables, X and Y, for which we can compute their

expected values, denoted as $\mathbb{E}[X]$ and $\mathbb{E}[Y]$. Let's assume that theoretically, we can establish that $\mathbb{E}[X] = \mathbb{E}[Y]$. Additionally, we can calculate their variances, denoted as $\text{Var}(X)$ and $\text{Var}(Y)$, and assume that theoretically, $\text{Var}(X) > \text{Var}(Y)$. Our objective is to estimate $\theta = \mathbb{E}[X]$, an unknown parameter, using Monte Carlo simulations by running n simulations with a standard error of $\frac{\sqrt{\text{Var}(X)}}{\sqrt{n}}$.

Notably, since we know that $\mathbb{E}[X] = \mathbb{E}[Y]$ and $\text{Var}(X) > \text{Var}(Y)$, we can observe that, with the same number of simulations using Monte Carlo technique on the random variable Y results in a narrower confidence interval compared to using the same method on X. Consequently, when estimating θ, it is more advantageous to utilize Y.

To exploit this observation and reduce variance, there are three primary variance reduction techniques available:

- **Antithetic Variables**: generating pairs of antithetic random variables, where each pair has opposite values. By averaging the results of these paired variables, you can reduce variance and obtain a more accurate estimate of θ.
- **Control Variables**: adding an extra variable in the simulation to decrease variance and enhance the estimation of θ.
- **Importance Sampling**: assigning varying weights to different simulation results according to their significance, thereby obtaining a lower-variance estimate of θ.

These variance reduction techniques aim to improve the efficiency of Monte Carlo simulations and provide more accurate estimates of the desired parameter θ. The technique that will be used in practice in the results in this Thesis is that of antithetic variables: let us explore it in detail.

We consider a framework that involves a random variable X and a function $g(X)$, where $g(X)$ is the price of a vanilla option. The aim is to estimate the expectation $\theta = \mathbb{E}[g(X)]$ and have the smallest possible variance. We can sample the mean as follows:

$$\hat{\theta}_n(X) = \frac{1}{n}\sum_{i=1}^{n} g\left(X_i\right)$$

The precision of this estimate can be quantified by its variance:

$$\text{Var}\left(\hat{\theta}_n\right) = \frac{\text{Var}(g(X))}{n} \quad \text{or} \quad \hat{\sigma}_n^2(g(X)) = \frac{1}{n-1}\sum_{i=1}^{n}\left(g\left(X_i\right) - \hat{\theta}_n(X)\right)^2$$

where $\hat{\sigma}_n^2(g(X))$ is the unbiased estimator of variance. The purpose is to reformulate the

problem in such a way that the new problem retains the same expected value but has a smaller variance. To achieve this, we introduce two random variables, X_1 and X_2, which share the same distribution as X. We define a new random variable, Y, as follows:

$$Y = \frac{1}{2}\left[g\left(X_1\right) + g\left(X_2\right)\right]$$

The expected value of Y doesn't change and it's:

$$\mathbb{E}[Y] = \frac{1}{2}\left[\mathbb{E}\left[g\left(X_1\right)\right] + \mathbb{E}\left[g\left(X_2\right)\right]\right] = \mathbb{E}[g(X)] \tag{3.10}$$

Hence, the new random variable Y has the same expected value as $g(X)$, but its variance is:

$$\begin{aligned} \operatorname{Var}(Y) &= \operatorname{Var}\left(\frac{g\left(X_1\right) + g\left(X_2\right)}{2}\right) \\ &= \frac{1}{4}\left[\operatorname{Var}\left(g\left(X_1\right)\right) + \operatorname{Var}\left(g\left(X_2\right)\right) + 2\operatorname{Cov}\left(g\left(X_1\right), g\left(X_2\right)\right)\right] \end{aligned} \tag{3.11}$$

Assuming that $\operatorname{Var}(g(X_1)) = \operatorname{Var}(g(X_2))$, we can simplify the variance of Y as:

$$\operatorname{Var}(Y) = \frac{1}{2}\left[\operatorname{Var}\left(g\left(X_1\right)\right) + \operatorname{Cov}\left(g\left(X_1\right), g\left(X_2\right)\right)\right] \tag{3.12}$$

If $g(X_1)$ and $g(X_2)$ are independent, the covariance term disappears, and we obtain the classical result for Monte Carlo simulation. However, if X_1 and X_2 are negatively correlated and $g(\cdot)$ is a monotone function, then the covariance term is negative. In this case, we have:

$$\operatorname{Var}(Y) = \operatorname{Var}\left(\frac{g\left(X_1\right) + g\left(X_2\right)}{2}\right) \leq \frac{1}{2}\operatorname{Var}\left(g\left(X_1\right)\right) \tag{3.13}$$

This demonstrates that by considering negatively correlated random variables X_1 and X_2 and exploiting the properties of $g(\cdot)$, we can achieve a reduced variance in the estimation of θ using the new random variable Y. To summarize, antithetic variables offer the advantage of smaller confidence intervals without a significant increase in computational effort compared to Monte Carlo simulations. The fundamental concept of antithetic variables consists in their negative correlation, as the name suggests. Whenever a random variable and its negatively correlated counterpart can be generated, the method becomes applicable. Generally, to obtain negative correlation between random variables uniformly distributed between 0 and 1, X_1 is taken randomly and $X_2 = 1 - X_1$ is defined such that they are both uniformly distributed and negatively correlated. In the specific case of using the standard normal distribution, we can set X_1 and $X_2 = -X_1$ so that both variables are normally distributed and negatively correlated.

3.2. Binomial tree method

In the previous section, a technique was introduced to be able to simulate the random variables contained in the underlying. In order to simulate different scenarios, random variables are sampled by knowing their probability distribution. In this section, however, we will introduce a second method that does not need to sample all the random components of the model, as we represent the set of paths of the underlying in a lattice. The technique used is that of the binomial tree (Cox, Ross and Rubinstein 1979, [10]), which is widely used in finance to price options, particularly vanilla options. A binomial tree, also known as a two-step model, is a mathematical approach used to break down the passage of time into distinct, discrete intervals when analyzing the pricing dynamics of an underlying asset. In this approach, it is presumed that the asset's value can either rise or fall during each time interval with well-defined probabilities. The term "binomial" is employed due to the fact that, during each interval, there are only two conceivable outcomes: an increase or a decrease in the asset's price. We will mainly analyze two methods of introducing the binomial tree model, one referring directly to the underlying, and the other related to the log-price.

3.2.1. The first step: price model

The CRR binomial tree model is based on the fact that the initial underlying is S_t has probability p to increase in value and probability one minus p to decrease in value. This event occurs at a given instant of time of the Δt $(= \frac{T}{n})$, calculated as the total time you want to consider for your analysis divided by the number of time steps you want to consider. As it is also intuitive to think, the greater the number of time steps is, the better the approximation to reality given by the model. In this model it is therefore important to specify the parameters, which are:

- u: growth coefficient of the variable S_t,
- d: coefficient of decreasing value of the variable S_t,
- p: probability that the value of the underlying will increase in value.

These parameters are closely related to those already "conventionally" present in the dynamics of S_t, namely volatility and interest rate.

As can be seen from Figure 3.1 the underlying has only two possibilities for the first time step; for subsequent time steps the situation remains similar.

Then looking at (Figure 3.1) and matching the mean and the variance of the S_t variable as

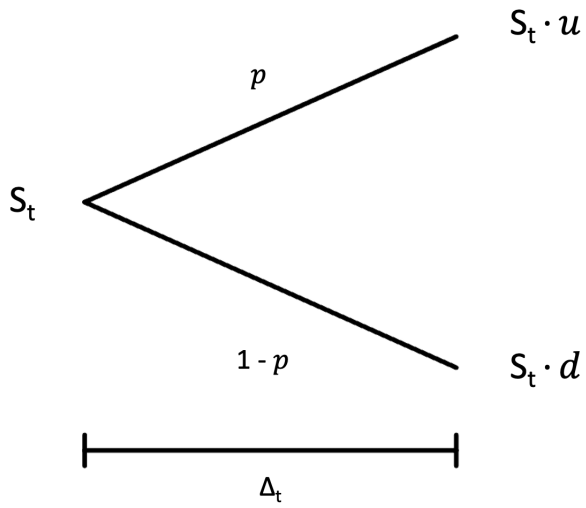


Figure 3.1: CRR binomial tree model: first step

a weighted probability combination of the two outcomes obtained, the following equations can be introduced:

$$\begin{cases} E\left[S_{t+\Delta t}\right] = S_t e^{r\Delta t} \\ \text{Var}\left(S_{t+\Delta t}\right) = E\left[S_{t+\Delta t}^2\right] - E\left[S_{t+\Delta t}\right]^2 \end{cases} \tag{3.14}$$

The first one is immediate:

$$p \cdot S_t u + (1-p) \cdot S_t d = E\left[S_{t+\Delta t}\right] = S_t e^{r\Delta t} \longrightarrow p = \frac{e^{r\Delta t} - d}{u - d} \tag{3.15}$$

The second is much more complex:

$$S_t^2 \cdot \sigma^2 \Delta t = S_t^2 \cdot [p \cdot u^2 + (1-p) \cdot d^2] - S_t^2 \cdot [p \cdot u + (1-p) \cdot d]^2 \tag{3.16}$$

so

$$\begin{aligned} \sigma^2 \Delta t &= p \cdot u^2 + (1-p) \cdot d^2 - p^2 \cdot u^2 - 2 \cdot p \cdot (1-p) \cdot u \cdot d - (1-p)^2 \cdot d^2 \\ &= p \cdot (1-p) \cdot \left[u^2 - 2 \cdot u \cdot d + d^2\right] = p \cdot (1-p) \cdot (u-d)^2 \end{aligned} \tag{3.17}$$

Let us focus on $p \cdot (1-p)$. We can write:

$$\begin{aligned} p \cdot (1-p) = p - p^2 &= \frac{e^{r\Delta t} - d}{u-d} - \frac{e^{2r\Delta t} - 2 \cdot d \cdot e^{r\Delta t} + d^2}{(u-d)^2} \\ &= \frac{e^{r\Delta t} \cdot u - u \cdot d - e^{r\Delta t} \cdot d + d^2 - e^{2r\Delta t} + 2 \cdot d \cdot e^{r\Delta t} - d^2}{(u-d)^2} \\ &= \frac{e^{r\Delta t}(u - d + 2 \cdot d) - u \cdot d - e^{2r\Delta t}}{(u-d)^2} = \frac{e^{r\Delta t}(u+d) - u \cdot d - e^{2r\Delta t}}{(u-d)^2} \end{aligned} \tag{3.18}$$

So now Eq. (3.17) can bhe seen as:

$$\sigma^2 \Delta t = e^{r\Delta t}(u+d) - u \cdot d - e^{2r\Delta t} \tag{3.19}$$

by defining $d = \frac{1}{u}$, we write:

$$\sigma^2 \Delta t = e^{r\Delta t}\left(u + \frac{1}{u}\right) - u \cdot \frac{1}{u} - e^{2r\Delta t} \longrightarrow u + \frac{1}{u} \quad = \frac{\sigma^2 \Delta t + 1 + e^{2r\Delta t}}{e^{r\Delta t}} \tag{3.20}$$

Considering that $e^{-r\Delta t} \approx (1 - r\Delta t), e^{r\Delta t} \approx (1 + r\Delta t),$ and $r \cdot \sigma^2 \cdot \Delta t^2 \to 0$, we gain

$$\begin{aligned} u + \frac{1}{u} &= e^{-r\Delta t}\sigma^2 \Delta t + e^{-r\Delta t} + e^{r\Delta t} \\ &\approx \sigma^2 \Delta t + 2 \end{aligned} \tag{3.21}$$

So

$$\begin{aligned} & u^2 - \left(\sigma^2 \Delta t + 2\right) u + 1 = 0 \\ & u = \frac{\sigma^2 \Delta t + 2 \pm \sqrt{\left(\sigma^2 \Delta t + 2\right)^2 - 4}}{2} \\ & \quad = \frac{\sigma^2 \Delta t + 2 \pm \sqrt{\sigma^4 \Delta t^2 + 4\sigma^2 \Delta t + 4 - 4}}{2} \quad \left(\sigma^4 \Delta t^2 \to 0\right) \\ & \quad = \frac{\sigma^2 \Delta t}{2} + 1 \pm \sigma\sqrt{\Delta t} \end{aligned} \tag{3.22}$$

Since $\sqrt{\Delta t}$ is much larger than Δt for a small Δt, and σ^2 is relatively smaller than σ, than we can ignore the first term $\frac{\sigma^2 \Delta t}{2}$ and we obtain:

$$u \approx 1 \pm \sigma\sqrt{\Delta t} \approx e^{+\sigma\sqrt{\Delta t}}$$

We consider only the value greater than 1 since u is a multiplicative growth coefficient.

By expanding for multiple time intervals as shown in Figure 3.2, the tree expands widely. Finally, it is possible to study the probability of states condition.

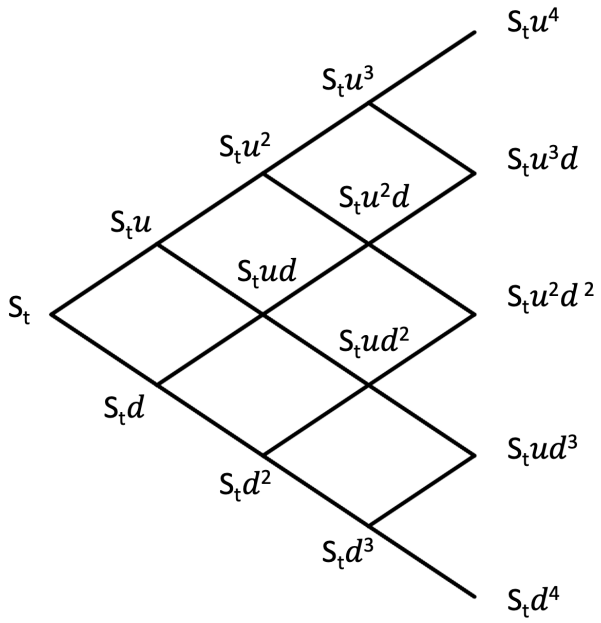


Figure 3.2: The general tree for the binomial model with n = 4

Defining the random variable with binomial distribution

$$Y = \begin{cases} u, \text{ with probability } p \\ d, \text{ with probability } 1-p \end{cases}$$

we can define $S_n = S_{n-1}Y_n, n \geq 1$. Given that the random multipliers are assumed to be i.i.d., the stochastic process $\{S_n : n \geq 0\}$ can be considered a Markov chain. This means that the probability of future states depends solely on the present state, conditioned on the past. The potential values of S_n are still determined by the probabilities following a binomial distribution with parameters n and p:

$$P\left(S_n = Su^k d^{n-k}\right) = \begin{pmatrix} n \\ k \end{pmatrix} p^k (1-p)^{n-k}, \quad 0 \leq k \leq n \tag{3.23}$$

3.2.2. The alternative model

Another binomial tree model introduced in Jarrow and Rudd's 1982 paper [29] involves utilizing the logarithmic representation of stock prices. Let us first recall the initial model of this Thesis in which the log-price dynamics was described by a drift coefficient μ and deterministic variance constant σ^2. So

$$X_t = \mu t + \sigma W_t, \quad t \geq 0, \tag{3.24}$$

and X_t is distributed as $N(\mu t, \sigma^2 t)$.
Thus, in our case, the value of the underlying asset is represented by a Geometric Brownian Motion:

$$S_t = S_0 e^{X_t} = S_0 e^{\mu t + \sigma W_t}, \quad t \geq 0 \tag{3.25}$$

Dividing by the constant $S(0)$ and taking logarithms, we get

$$\ln(S_t/S_0) = X_t = \mu t + \sigma W_t, \quad t \geq 0. \tag{3.26}$$

Notice that S_t has a lognormal distribution. The reasoning applied is similar to that in the previous section, but reference is made to the log-price.

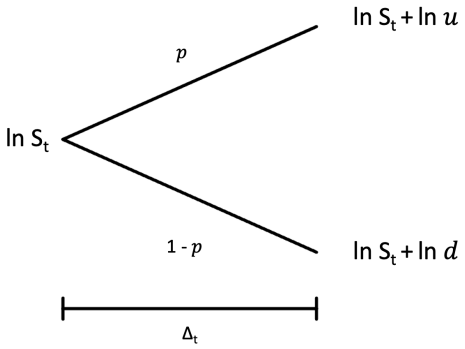


Figure 3.3: Binomial tree by Jarrow and Rudd: first step

U and D, shown in Figure 3.3, are defined as $ln(u)$ and $-ln(u)$ $(= ln(d))$ respectively, and the mean and variance of the lognormal distribution are reasoned:

$$\begin{cases} E\left[\ln\left(S_n/S_0\right)\right] = nE\left[\ln\left(S_1/S_0\right)\right] = n\mu\Delta t \\ \mathrm{Var}\left(\ln\left(S_n/S_0\right)\right) = n\,\mathrm{Var}\left(\ln\left(S_1/S_0\right)\right) = n\sigma^2\Delta t \end{cases} \tag{3.27}$$

also seen as

$$\begin{cases} E\left[\ln\left(S_1/S_0\right)\right] = \mu\Delta t \\ \mathrm{Var}\left(\ln\left(S_1/S_0\right)\right) = \sigma^2\Delta t \end{cases} \tag{3.28}$$

thanks to the fact that $\{X_n, n > 0\}$ are i.i.d. Again, the first equation is resolvable rapidly; we rewrite:

$$pU + (1-p)(-U) = (2p-1)U = \mu\Delta t \longrightarrow p = \frac{\mu\Delta t}{2U} + \frac{1}{2} \tag{3.29}$$

while, for computational simplicity, for the second term we focus on the moment second

instead of the variance:

$$pU^2 + (1-p)(-U)^2 = U^2 = \sigma^2\Delta t + (\mu\Delta t)^2 \tag{3.30}$$

These equations can be combined to form a single quadratic equation. However, a less formal simplification is made than in the previous section: Δt should tend to be as small as possible, and thus Δt^2 will result in even smaller. We can deduce that $(\mu\Delta t)^2$ should be negligible compared to $\sigma^2\Delta t$. If indeed $(\mu\Delta t)^2$ is suitably small relative to $\sigma^2\Delta t$, we can safely omit the term $(\mu\Delta t)^2$ with minimal loss of precision:

$$U^2 \approx \sigma^2\Delta t \tag{3.31}$$

which immediately yields to

$$U = \sigma\sqrt{\Delta t} \text{ and } u = e^U = e^{\sigma\sqrt{\Delta t}} \tag{3.32}$$

We remember that $d = 1/u$, we also have

$$d = e^{-U} = e^{-\sigma\sqrt{\Delta t}} \tag{3.33}$$

Finally, Combining Eq. (3.29) and Eq. (3.32), we obtain

$$p = \frac{1}{2} + \frac{\mu}{2\sigma}\sqrt{\Delta t} \tag{3.34}$$

Although the mean remains accurate, there is a small error in the second moment and the variance.

3.2.3. Derivatives with binomial tree

The calculation of options using this method can be divided mainly into two categories: one is referred to options in which there is no possibility of exercising before maturity and in which the path of the underlying in the time interval considered does not affect the value of the derivative, such as classic European Options. The other, instead, refers to options in which the path is fundamental to the pricing and/or in which there is the possibility of exercising the option before it expires, such as American Options:

- The first category of options is much simpler to compute. In particular, in this context, after constructing the binomial tree and thus obtaining the value of the underlying at the final time instant, it is possible to simply proceed to compute the

vector of pay-offs, depending on the derivative we are going to consider, and then compute the expected value and discount it using the interest-rate.

- The second category, is slightly more complicated. The key to valuing an option with the binomial model is the concept of "backward induction". We start with the maturity of the option and calculate its value at that time; next, we backward value the option. To estimate the option price at each node in the tree, we need to consider two possibilities for each time step: the underlying asset may increase or decrease by a certain amount. If the option is profitable, its intrinsic value, which is the difference between the price of the underlying asset and the strike price, is positive; otherwise, it is zero.This back-valuation process is repeated until the present time is reached, providing the option's present price. This price represents what the option is worth in present terms, taking into account future price fluctuations in the underlying asset.

The result introduced in the next chapter, since we are dealing with European call options, refer to the first of the two categories.

3.2.4. Binomial tree for a Wiener process

In the previous section, deterministic variance made the model simple, resulting in a recombinant binomial tree. This means that starting at any node and having a growth followed by a decrease in the underlying or vice versa, leads to the same result.

Given the complexity introduced in Chapter 2, the tree constructed (see Section 3.3.4) does not turn out to be recombinant. In this section we want to build only the dynamics of Wiener process (W_t in Eq. (3.35)) that is fundamental for subsequent calculations. W_t willl be recombinant too: starting from the dynamics of S_t, simplifying its mean and variance to obtain the dynamics of e^{W_t} and finally the one of W_t. Setting $\mu = 0$ and $\sigma = 1$, Eq. (3.35) becomes

$$X_t = W_t, \quad t \geq 0, \tag{3.35}$$

We can reconsider d, u and p as

$$d = e^{-\sigma\sqrt{\Delta t}}$$
$$u = e^{\sigma\sqrt{\Delta t}}$$
$$p = \frac{1}{2}$$

Recalling that W_0 is equal to 0, it is possible to recreate the tree for e^{W_t}, as shown in

Figure 3.4:

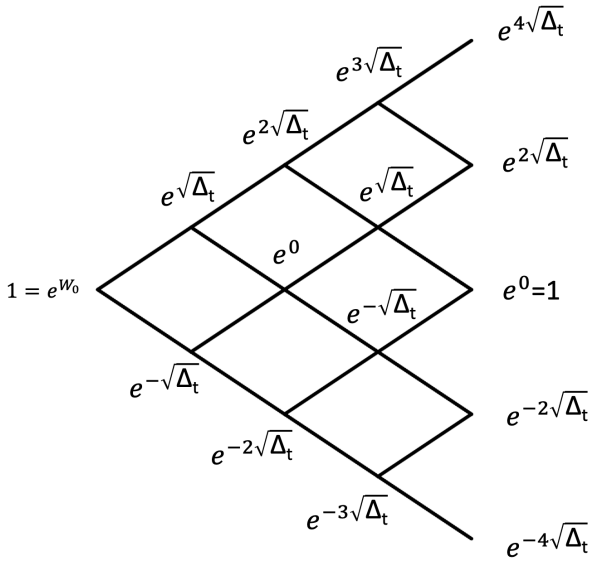


Figure 3.4: The binomial tree for a exponential Wiener process (GBM) with n = 4

As anticipated this is again a recombinant tree. Our purpose, however, is to obtain the structure for the W_t process only and not e^{W_t}. We therefore proceed by eliminating the exponential function as shown in Figure 3.5. Also in this context $p = \frac{1}{2}$:

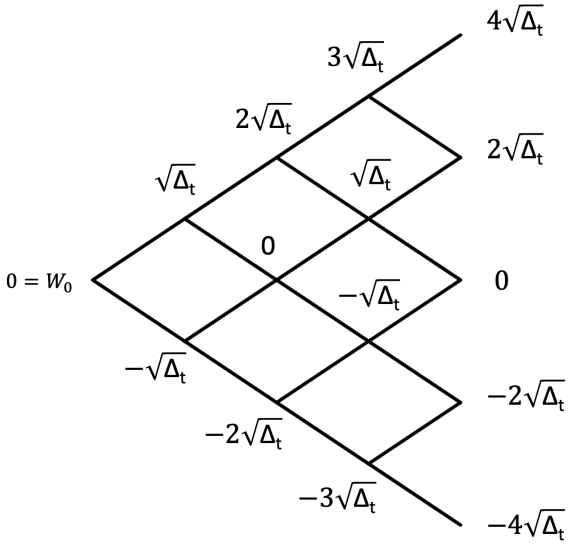


Figure 3.5: The binomial tree for a Wiener (BM) process with n = 4

3.2.5. Intuitive feedback

The results obtained in the Subsection 3.2.4 coincide with what can be intuitively thought. In fact, it is enough to consider that the W_t (Wiener process) is itself a Brown motion, with zero mean and variance equal to Δt. Since the distribution is normal, and since W_t has zero initial value $W_0 = 0$, it is obvious that the probability of increasing or decreasing of the same quantity are identical ($p = \frac{1}{2}$). Starting from the dynamics of this process and lowering it into the equations of Path-Dependent Volatility models, we are able to construct a non-recombinant binomial tree.

3.3. Time step discretization

In this section we will look at techniques by which we can discretize our model over time to apply both simulation methods. The discretizations used are the main ones in the world of finance: Euler scheme and Milstein scheme. Although these techniques are well documented, their applications vary for each model. In this Thesis, we have developed the appropriate calculations to make both pricing techniques applicable to the PDV model.

3.3.1. Euler scheme for Monte Carlo: 2 and 4-factor models

Let us consider the following situation where

$$S_t = S_0 e^{X_t} \quad \text{and} \quad X_t = \mu(S_t)t + \sigma(S_t)W_t, \quad t \geq 0 \tag{3.36}$$

Starting from the differential form of the underlying and considering that in our models of Chapter 2, $\mu(S_t)$ is always equal to zero, we have

$$\begin{cases} dS_t = \sigma(S_t)S_t dW_t \\ S_0 \end{cases} \tag{3.37}$$

and applying Itô's formula, it can be derived the differential form for X_t

$$\begin{cases} dln(S_t) = dX_t = -\dfrac{\sigma(S_t)^2}{2}t + \sigma(S_t)dW_t \\ X_0 \end{cases} \tag{3.38}$$

Moving to the integral form, with Δt small, we can rewrite X_t as

$$X_{t+\Delta_t} = X_t + \int_t^{t+\Delta t} -\frac{\sigma\left(S_u\right)^2}{2}du + \int_t^{t+\Delta t} \sigma\left(S_u\right)dW_u \tag{3.39}$$

where $t < s < t + \Delta_t$. The approximations that characterize Euler's scheme [2] are:

$$\int_t^{t+\Delta t} -\frac{\sigma\left(S_u\right)^2}{2} du \approx -\frac{\sigma\left(S_t\right)^2}{2}\Delta t \tag{3.40}$$

$$\int_t^{t+\Delta t} \sigma\left(S_u\right) dW_u \approx \sigma\left(S_t\right)\left[W_{t+\Delta t} - W_t\right] \tag{3.41}$$

The first, in a very intuitive way, holds constant at the initial time of the integral t the value of the argument, multiplying it by only Δ_t, since this value is considered to be very small. The second approximation, slightly more complex since it involves stochastic integrals, keeps yes constant the argument of the integral at time t, which is, however, multiplied by the difference of two Wiener processes. By definition, that difference, is a random variable distributed as a normal of zero mean and variance Δ_t, so

$$\int_t^{t+\Delta t} \sigma\left(S_u\right) dW_u \approx \sigma\left(S_t\right)\sqrt{\Delta_t} Z_{t+\Delta t} \tag{3.42}$$

where $Z_{t+\Delta t} \sim \mathcal{N}\left(0,1\right)$ evaluated at the time $t + \Delta t$. Applying this reasoning to PDV 2-factor model (2.10) wherever a stochastic integral is present, the following discretized model is obtained:

$$\begin{aligned}
X_{t+\Delta t} &= X_t + \left(-\frac{1}{2}\sigma_t^2\right)\Delta t + \sigma_t\sqrt{\Delta t} Z_{t+\Delta t} \\
\sigma_t &= \sigma\left(R_{1,t}, R_{2,t}\right) = \beta_0 + \beta_1 R_{1,t} + \beta_2\sqrt{R_{2,t}} \\
R_{1,t+\Delta t} &= R_{1,t} + \lambda_1\left(\sigma\left(R_{1,t}, R_{2,t}\right)\sqrt{\Delta t} Z_{t+\Delta t} - R_{1,t}\Delta t\right) \\
R_{2,t+\Delta t} &= R_{2,t} + \lambda_2\left(\sigma\left(R_{1,t}, R_{2,t}\right)^2 - R_{2,t}\right)\Delta t
\end{aligned} \tag{3.43}$$

Similarly, it is possible to obtain the concretized equivalent of the 4-factor model (2.17):

$$\begin{aligned}
X_{t+\Delta t} &= X_t + \left(-\frac{1}{2}\sigma_t^2\right)\Delta t + \sigma_t\sqrt{\Delta t} Z_{t+\Delta t} \\
\sigma_t &= \sigma\left(R_{1,t}, R_{2,t}\right) = \beta_0 + \beta_1 R_{1,t} + \beta_2\sqrt{R_{2,t}} \\
R_{1,t} &= \left(1-\theta_1\right) R_{1,0,t} + \theta_1 R_{1,1,t} \\
R_{2,t} &= \left(1-\theta_2\right) R_{2,0,t} + \theta_2 R_{2,1,t} \\
R_{1,j,t+\Delta t} &= R_{1,j,t} + \lambda_{1,j}\left(\sigma\left(R_{1,t}, R_{2,t}\right)\sqrt{\Delta t} Z_{t+\Delta t} - R_{1,j,t}\Delta t\right), \quad j \in \{0,1\} \\
R_{2,j,t+\Delta t} &= R_{2,j,t} + \lambda_{2,j}\left(\sigma\left(R_{1,t}, R_{2,t}\right)^2 - R_{2,j,t}\right)\Delta t, \quad j \in \{0,1\}
\end{aligned} \tag{3.44}$$

In both cases X_t is to be considered the log-price and Z_t the variable to be sampled to apply the Monte Carlo method.

3.3.2. Milstein scheme in Monte Carlo: 2-factor model

The key concept behind the Milstein scheme [30, 31] is that it enhances the accuracy of discretization by considering expansions of the coefficients $\mu(S_t)$ and $\sigma(S_t)$ via Itô's lemma, since the coefficients are functions of S_t. As before, we consider $\mu(S_t) = 0$. The SDE for the coefficient is:

$$d\sigma_t = \left(\sigma_t'\mu_t + \frac{1}{2}\sigma_t''\sigma_t^2\right) dt + (\sigma_t'\sigma_t)\, dW_t \tag{3.45}$$

In the context of this expression, the prime notation denotes differentiation with respect to S_t, and it's important to note that the derivatives with respect to t are assumed to be zero due to our assumption that σ_t does not directly depend on t. Let's now focus on the 2-factor model: as already mentioned in Eq. (3.48), the differential of σ can be seen as

$$d\sigma_t = \left(-\beta_1\lambda_1 R_{1,t} + \frac{\beta_2\lambda_2}{2}\frac{\sigma_t^2 - R_{2,t}}{\sqrt{R_{2,t}}}\right) dt + \beta_1\lambda_1\sigma_t dW_t \tag{3.46}$$

Consequently, the coefficients at time are expressed in integral form as follows:

$$\sigma_{t+\Delta t} = \sigma_t + \int_t^{t+\Delta t}\left(\sigma_s' + \frac{1}{2}\sigma_s''\sigma_s^2\right) ds + \int_t^{t+\Delta t}(\sigma_s'\sigma_s)\, dW_s \tag{3.47}$$

or as

$$\sigma_{t+\Delta t} = \sigma_t + \int_t^{t+\Delta t}\left(-\beta_1\lambda_1 R_{1,s} + \frac{\beta_2\lambda_2}{2}\frac{\sigma_s^2 - R_{2,s}}{\sqrt{R_{2,s}}}\right) ds + \int_t^{t+\Delta t}\beta_1\lambda_1\sigma_s dW_s \tag{3.48}$$

Reconsidering the equation to discretize (3.39), we can proceed as before, but considering the expansion of σ_t. The terms higher than order one as $ds{\cdot}du = O\left((dt)^2\right)$ and $ds{\cdot}dWu = O\left((dt)^{3/2}\right)$ can be ignored given their infinitesimal contribution. First we calculate the approximation for the integral in time:

$$\int_t^{t+\Delta t} -\frac{\sigma\left(S_u\right)^2}{2} du \approx -\frac{\sigma\left(S_t\right)^2}{2}\Delta t \tag{3.49}$$

Nothing has changed since the σ^2 is the sum of σ, σds and σdWs and, composed with du in the integal of (3.39), only σ^2 remains, so the computation is as before. In contrast,

the second integral, namely the stochastic integral, is more complex:

$$\begin{aligned}\int_t^{t+\Delta t}\sigma\left(S_u\right)dW_u &= \int_t^{t+\Delta t}\left(\sigma_t+\int_t^u\left(-\beta_1\lambda_1R_{1,s}+\frac{\beta_2\lambda_2}{2}\frac{\sigma_s^2-R_{2,s}}{\sqrt{R_{2,s}}}\right)ds\right.\\ &\left.+\int_t^u\beta_1\lambda_1\sigma_s dW_s\right)dW_u\\ &=\int_t^{t+\Delta t}\sigma_t dW_u\\ &+\int_t^{t+\Delta t}\left(\int_t^u\left(-\beta_1\lambda_1R_{1,s}+\frac{\beta_2\lambda_2}{2}\frac{\sigma_s^2-R_{2,s}}{\sqrt{R_{2,s}}}\right)ds\right)dW_u\\ &+\int_t^{t+\Delta t}\left(\int_t^u\beta_1\lambda_1\sigma_s dW_s\right)dW_u\\ &\approx\int_t^{t+\Delta t}\sigma_t dW_u+\int_t^{t+\Delta t}\left(\int_t^u\beta_1\lambda_1\sigma_s dW_s\right)dW_u,\end{aligned} \tag{3.50}$$

recalling that the product $ds\cdot dWu$ leads to negligible terms. The first integral, as before, is

$$\begin{aligned}\int_t^{t+\Delta t}\sigma_t dW_u &\approx \sigma\left(S_t\right)\left[W_{t+\Delta t}-W_t\right]\\ &\approx \sigma_t\sqrt{\Delta_t}Z_{t+\Delta t}\end{aligned} \tag{3.51}$$

The second and more complex is:

$$\begin{aligned}\int_t^{t+\Delta t}\left(\int_t^u\beta_1\lambda_1\sigma_s dW_s\right)dW_u &\approx \beta_1\lambda_1\sigma_t\int_t^{t+\Delta t}\int_t^u dW_s\cdot dW_u\\ &=\beta_1\lambda_1\sigma_t\int_t^{t+\Delta t}\left(W_u-W_t\right)\cdot dW_u\\ &=\beta_1\lambda_1\sigma_t\left(\int_t^{t+\Delta t}W_s dW_s-W_tW_{t+\Delta t}+{W_t}^2\right)\end{aligned} \tag{3.52}$$

In order to solve the last integral we introduce a new variable Y_t. Using Itô's Lemma, it can be easily shown that Y_t has a solution[2] $Y_t=\frac{1}{2}W_t^2-\frac{1}{2}t$ such that

$$\int_t^{t+\Delta t}W_s dW_s = Y_{t+\Delta t}-Y_t=\frac{1}{2}W_{t+\Delta t}^2-\frac{1}{2}W_t^2-\frac{1}{2}\Delta t \tag{3.53}$$

[2] Indeed, $\frac{\partial Y}{\partial t}=-\frac{1}{2}$, $\frac{\partial Y}{\partial W}=W$, and $\frac{\partial^2 Y}{\partial W^2}=1$, so that $dY_t=\left(-\frac{1}{2}+0+\frac{1}{2}\cdot 1\cdot 1\right)dt+\left(W_t\cdot 1\right)dW_t=W_t dW_t$.

Substitute back into (3.52) to obtain

$$\int_t^{t+\Delta t}\left(\int_t^u \beta_1\lambda_1\sigma_s dW_s\right)dW_u \approx \frac{1}{2}\beta_1\lambda_1\sigma_t\left[(W_{t+\Delta t}-W_t)^2-\Delta t\right]$$
$$= \frac{1}{2}\beta_1\lambda_1\sigma_t\left[(\Delta W_t)^2-\Delta t\right]$$

where $W_{\Delta t}=W_{t+\Delta t}-W_t$, which is equal in distribution to $\sqrt{\Delta t}Z$ with Z distributed as standard normal. Similar reasoning should be carried out on R_1 since it too, an integral form, has a stochastic integral that can thus be expanded more than Euler's scheme We can now rewrite the 2-factor model in discrete useful form using the Milstein scheme:

$$\begin{aligned}
X_{t+\Delta t} &= X_t + \left(-\frac{1}{2}\sigma_t^2\right)\Delta t + \sigma_t\sqrt{\Delta t}Z_{t+\Delta t} + \frac{1}{2}\lambda_1\beta_1\sigma_t\left(Z_{t+\Delta t}^2-1\right)\Delta t \\
\sigma_t &= \sigma\left(R_{1,t},R_{2,t}\right) = \beta_0+\beta_1 R_{1,t}+\beta_2\sqrt{R_{2,t}} \\
R_{1,t+\Delta t} &= R_{1,t}+\lambda_1\left(\sigma\left(R_{1,t},R_{2,t}\right)\sqrt{\Delta t}Z_{t+\Delta t}-R_{1,t}\Delta t\right)+\frac{1}{2}\lambda_1^2\beta_1\sigma_t\left(Z_{t+\Delta t}^2-1\right)\Delta t \\
R_{2,t+\Delta t} &= R_{2,t}+\lambda_2\left(\sigma_t\left(R_{1,t},R_{2,t}\right)^2-R_{2,t}\right)\Delta t
\end{aligned} \tag{3.54}$$

3.3.3. Milstein scheme in Monte Carlo: 4-factor model

In order to apply the Milstein scheme to the 4-factor model, we need to go over what was demonstrated in the previous section; in particular the differences come in the calculation of the stochastic integral (3.52). We begin by introducing the differential form of σ_t the for this model:

$$d\sigma_t = \left(-\beta_1\bar{\lambda}_1\bar{R}_{1,t}+\frac{\beta_2\bar{\lambda}_2}{2}\frac{\sigma_t^2-\bar{R}_{2,t}}{\sqrt{R_{2,t}}}\right)dt+\beta_1\bar{\lambda}_1\sigma_t dW_t \tag{3.55}$$

where

$$\bar{\lambda}_j := (1-\theta_j)\lambda_{j,0}+\theta_j\lambda_{j,1}, \quad j\in\{0,1\}$$
$$\bar{R}_{j,t} := \frac{(1-\theta_j)\lambda_{j,0}R_{j,0,t}+\theta_j\lambda_{j,1}R_{j,1,t}}{\bar{\lambda}_j}, \quad j\in\{0,1\}$$

as introduced in Eq. (2.16). In this case we have:

$$\int_t^{t+\Delta t}\sigma\left(S_u\right)dW_u \approx \sigma_t\sqrt{\Delta_t}Z_{t+\Delta t}+\frac{1}{2}\bar{\lambda}_1\beta_1\sigma_t\left(Z_{t+\Delta t}^2-1\right)\Delta t \tag{3.56}$$

Finally, the dynamics is

$$\begin{aligned}
X_{t+\Delta t} &= X_t + \left(-\frac{1}{2}\sigma_t^2\right)\Delta t + \sigma_t\sqrt{\Delta t}Z_{t+\Delta t} + \frac{1}{2}\bar{\lambda}_1\beta_1\sigma_t\left(Z_{t+\Delta t}^2 - 1\right)\Delta t \\
\sigma_t &= \sigma\left(R_{1,t}, R_{2,t}\right) = \beta_0 + \beta_1 R_{1,t} + \beta_2\sqrt{R_{2,t}} \\
R_{1,t} &= \left(1-\theta_1\right)R_{1,0,t} + \theta_1 R_{1,1,t} \\
R_{2,t} &= \left(1-\theta_2\right)R_{2,0,t} + \theta_2 R_{2,1,t} \\
R_{1,j,t+\Delta t} &= R_{1,j,t} + \lambda_{1,j}\left(\sigma_t\sqrt{\Delta t}Z_{t+\Delta t} - R_{1,j,t}\Delta t\right) + \frac{1}{2}\bar{\lambda}_1\lambda_{1,j}\beta_1\sigma_t\left(Z_{t+\Delta t}^2 - 1\right)\Delta t \\
R_{2,j,t+\Delta t} &= R_{2,j,t} + \lambda_{2,j}\left(\sigma_t^2 - R_{2,j,t}\right)\Delta t
\end{aligned} \tag{3.57}$$

where, $j \in \{0,1\}$. Again, we have to keep in mind that the dynamics of R_1 have been expanded and that X_t is always the log-price.

3.3.4. Binomial Tree discretization

We have seen how discretization techniques can be applied with the Monte Carlo method, but the issue is slightly different for the binomial tree. Let us consider the 2-factor model and recover the results proposed for Euler scheme with respect to the approximation of the integral of Eq. (3.39). In particular, we consider the approximation (3.41):

$$\int_t^{t+\Delta t} \sigma\left(S_u\right) dW_u \approx \sigma\left(S_t\right)\left[W_{t+\Delta t} - W_t\right]$$

In this approximation the dynamics of $[W_{t+\Delta t} - W_t]$ can be represented as a binomial tree as shown in fFe that, since the dynamics of R_1 also includes a random factor, we need to carry over time all the dynamics of the 2-factor model (R_1, R_2, X_t, and σ_t) simultaneously. It is also immediate to note that the dynamics that compose the model are not recombinant.

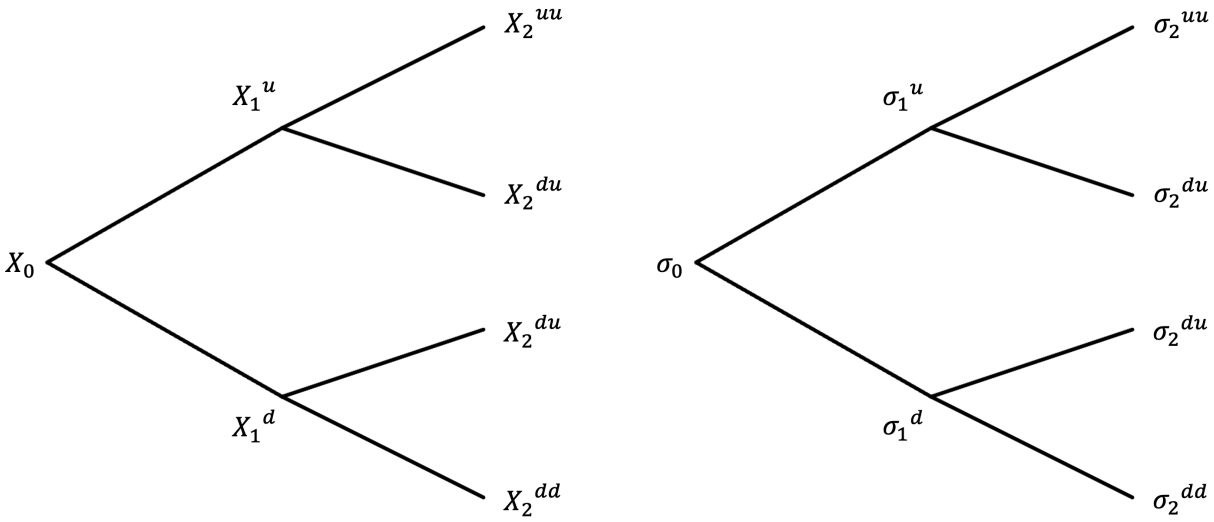


Figure 3.6: Binomial tree for X_t and σ_t in the 2-factor model

Figure 3.6, although the time interval considered are only two, shows how much the binomial tree of X and sigma processes can expand. In particular, the number of values at a given instant of time is exponential with respect to n, and the same holds for process R_1 and R_2 in Figure 3.7. All initial values are calculated by calibrating the model to empirical data.

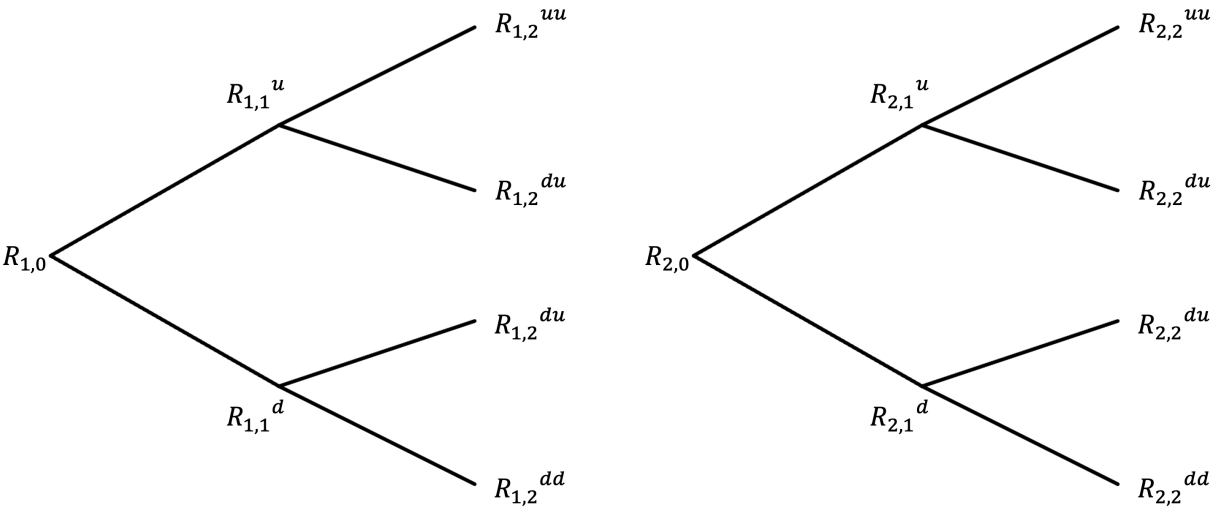


Figure 3.7: Binomial tree for R_1 and R_2 in the 2-factor model

The idea behind applying the binomial model of the Wiener process to all dynamics of the 2-factor model is to replace, at each time interval, the stochastic component, denoted

by $\Delta t Z$ in the previous section, with the same possibility of increasing or decreasing $\sqrt{\Delta t}$, as shown $W_{\Delta t}$. We will use the following discretization with nontrivial expedients

$$\begin{aligned} X_{t+\Delta t} &= X_t + \left(-\frac{1}{2}\sigma_t^2\right)\Delta t + \sigma_t W_{\Delta t} \\ \sigma_t &= \sigma\left(R_{1,t}, R_{2,t}\right) = \beta_0 + \beta_1 R_{1,t} + \beta_2\sqrt{R_{2,t}} \\ R_{1,t+\Delta t} &= R_{1,t} + \lambda_1\left(\sigma\left(R_{1,t}, R_{2,t}\right) W_{\Delta t} - R_{1,t}\Delta t\right) \\ R_{2,t+\Delta t} &= R_{2,t} + \lambda_2\left(\sigma\left(R_{1,t}, R_{2,t}\right)^2 - R_{2,t}\right)\Delta t \end{aligned} \tag{3.58}$$

If for Monte Carlo it was enough simply to replace the Wiener process with a sampling of the normal distribution, for the binomial tree the path must be considered.

We can therefore develop the first time step for X_t:

$$\begin{aligned} X_1^u &= X_0 + \left(-\frac{1}{2}\sigma_0^2\right)\Delta t + \sigma_0[\sqrt{\Delta t}] \\ X_1^d &= X_0 + \left(-\frac{1}{2}\sigma_0^2\right)\Delta t + \sigma_0[-\sqrt{\Delta t}] \end{aligned}$$

The first step in the binomial model for the Wiener process has been introduced.

Considering this, we can define the first step for all the dinamics:

- $R_{1,1}$:

$$\begin{aligned} R_{1,1}^u &= R_{1,0} + \lambda_1\left(\sigma_0[\sqrt{\Delta t}] - R_{1,0}\Delta t\right) \\ R_{1,1}^d &= R_{1,0} + \lambda_1\left(\sigma_0[-\sqrt{\Delta t}] - R_{1,0}\Delta t\right) \end{aligned}$$

- $R_{2,1}$:

$$\begin{aligned} R_{2,1}^u &= R_{2,0} + \lambda_2\left(\sigma_0^2 - R_{2,0}\right)\Delta t \\ R_{2,1}^d &= R_{2,0} + \lambda_2\left(\sigma_0^2 - R_{2,0}\right)\Delta t \end{aligned}$$

- σ_1:

$$\begin{aligned} \sigma_1^u &= \beta_0 + \beta_1 R_{1,1}^u + \beta_2\sqrt{R_{2,1}^u} \\ \sigma_1^d &= \beta_0 + \beta_1 R_{1,1}^d + \beta_2\sqrt{R_{2,1}^d} \end{aligned}$$

In this first step we begin to understand the complexity and precision that is required to proceed along the tree. Consider the variable σ: this is composed of the values of the dynamics of R_1 and R_2, which can increase or decrease. To calculate the increased amount of σ, we must consider the increased values of R_1 and R_2. Similarly when it decreases.

This concept becomes clearer as early as the following step:

- X_2:

$$X_1^{uu} = X_1^u + \left(-\frac{1}{2}\sigma_1^{u2}\right)\Delta t + \sigma_1^{u2}[\sqrt{\Delta t}]$$
$$X_1^{ud} = X_1^u + \left(-\frac{1}{2}\sigma_1^{u2}\right)\Delta t + \sigma_1^{u2}[-\sqrt{\Delta t}]$$
$$X_1^{du} = X_1^d + \left(-\frac{1}{2}\sigma_1^{d2}\right)\Delta t + \sigma_1^{d2}[\sqrt{\Delta t}]$$
$$X_1^{dd} = X_1^d + \left(-\frac{1}{2}\sigma_1^{d2}\right)\Delta t + \sigma_1^{d2}[-\sqrt{\Delta t}]$$

- $R_{1,2}$:

$$R_{1,2}^{uu} = R_{1,1}^u + \lambda_1\left(\sigma_1^u[\sqrt{\Delta t}] - R_{1,1}^u\Delta t\right)$$
$$R_{1,2}^{ud} = R_{1,1}^u + \lambda_1\left(\sigma_1^u[-\sqrt{\Delta t}] - R_{1,1}^u\Delta t\right)$$
$$R_{1,2}^{du} = R_{1,1}^d + \lambda_1\left(\sigma_1^d[\sqrt{\Delta t}] - R_{1,1}^d\Delta t\right)$$
$$R_{1,2}^{dd} = R_{1,1}^d + \lambda_1\left(\sigma_1^d[-\sqrt{\Delta t}] - R_{1,1}^d\Delta t\right)$$

- $R_{2,2}$:

$$R_{2,2}^{uu} = R_{2,1}^u + \lambda_2\left(\sigma_1^{u2} - R_{2,1}^u\right)\Delta t$$
$$R_{2,2}^{ud} = R_{2,1}^u + \lambda_2\left(\sigma_1^{u2} - R_{2,1}^u\right)\Delta t$$
$$R_{2,2}^{du} = R_{2,1}^d + \lambda_2\left(\sigma_1^{d2} - R_{2,1}^d\right)\Delta t$$
$$R_{2,2}^{dd} = R_{2,1}^d + \lambda_2\left(\sigma_1^{d2} - R_{2,1}^d\right)\Delta t$$

- σ_2:

$$\sigma_2^{uu} = \beta_0 + \beta_1 R_{1,2}^{uu} + \beta_2\sqrt{R_{2,2}^{uu}}$$
$$\sigma_2^{ud} = \beta_0 + \beta_1 R_{1,2}^{ud} + \beta_2\sqrt{R_{2,2}^{uu}}$$
$$\sigma_2^{du} = \beta_0 + \beta_1 R_{1,2}^{du} + \beta_2\sqrt{R_{2,2}^{du}}$$
$$\sigma_2^{dd} = \beta_0 + \beta_1 R_{1,2}^{dd} + \beta_2\sqrt{R_{2,2}^{dd}}$$

Consider, for example, the value X_2^{uu}, which depends on σ_1. In this particular case it is not possible to consider σ_1^d, the value of sigma that at time one had a decrease in value. In fact, the first u of X_2^{uu} indicates that this path at time one went up, and this

is true for all the dynamics that make up that value, so we must consider σ_1^u. All this is caused by the fact that the tree does not turn out to be recombinant. The same reasoning applies when any variable travels along a path on which it decreases. This reasoning must be kept in mind for all dynamics and all time instants, and this effect from simple discretization (3.58) is not known.

In a similar way, it is possible to proceed with the 4-factor model:

$$\begin{aligned}
X_{t+\Delta t} &= X_t + \left(-\frac{1}{2}\sigma_t^2\right)\Delta t + \sigma_t W_{\Delta t} \\
\sigma_t &= \sigma\left(R_{1,t}, R_{2,t}\right) = \beta_0 + \beta_1 R_{1,t} + \beta_2\sqrt{R_{2,t}} \\
R_{1,t} &= \left(1-\theta_1\right)R_{1,0,t} + \theta_1 R_{1,1,t} \\
R_{2,t} &= \left(1-\theta_2\right)R_{2,0,t} + \theta_2 R_{2,1,t} \\
R_{1,j,t+\Delta t} &= R_{1,j,t} + \lambda_{1,j}\left(\sigma\left(R_{1,t}, R_{2,t}\right)W_{\Delta t} - R_{1,j,t}\Delta t\right), \quad j \in \{0,1\} \\
R_{2,j,t+\Delta t} &= R_{2,j,t} + \lambda_{2,j}\left(\sigma\left(R_{1,t}, R_{2,t}\right)^2 - R_{2,j,t}\right)\Delta t, \quad j \in \{0,1\}
\end{aligned} \tag{3.59}$$

with proper consideration in the development.

Finally we see the non-applicability of the Milstein scheme discretization to the binomial tree for this model. The fundamental approximation, again in the case of the 2-factor model, is

$$\int_t^{t+\Delta t} \sigma\left(S_u\right) dW_u \approx \sigma_t W_{\Delta t} + \frac{1}{2}\lambda_1\beta_1\sigma_t\left(W_{\Delta t}^2 - \Delta t\right) \tag{3.60}$$

Considering the fact that the binomial tree of $\sigma_t W_{\Delta t}$ at each time interval can increase or decrease by $\sqrt{\Delta t}$, it results that

$$\left([\pm\sqrt{\Delta t}]^2 - \Delta t\right) = 0 \tag{3.61}$$

and thus null the additional component of the Milstein scheme. The same holds true for the 4-factor model.

3.4. VIX: implementation

In 2004, the CBOE Futures Exchange (CFE) began trading VIX futures, which were calculated based on a portfolio of options, as shown in papers [16] and [17]. The VIX index is a measure of market volatility that is constructed from options, rather than stocks; in particular the VIX Index measures 30-day expected volatility of the S&P 500 Indexand the calculation takes as input the market prices of SPX options. Its price reflects

the market's anticipation of future volatility and similar to traditional indices, the VIX index can be computed using a formula:

$$\left(\frac{VIX}{100}\right)^2 = \frac{2}{T}\sum_i \frac{\Delta K_i}{K_i^2} e^{RT} Q\left(K_i\right) - \frac{1}{T}\left[\frac{F}{K_0} - 1\right]^2 \tag{3.62}$$

where we have:

- T Time to expiration
- F Forward index level derived from index option prices
- K_0 First strike below the forward index level, F
- K_i Strike price of i^{it} out-of-the-money option; a call if $K_i > K_0$ and a put if $K_i < K_0$; both put and call if $K_i = K_0$
- $\Delta K_i = \frac{K_{i+1} - K_{i-1}}{2}$ (Interval between strike prices - half of the difference between strike prices K_{i-1} and K_{i+1})
- R Risk-free interest rate to expiration
- $Q\left(K_i\right)$ The midpoint of the bid-ask spread for each option with strike K_i

Since this model turns out to be very complex, alternative ways must be found.

3.4.1. VIX definition

Most of the current literature on volatility derivatives using a stochastic volatility model as the basic concept: this approach establishes a connection between the options market and the volatility derivatives market, but considers the correlation between the variance and the pricing process (of the S&P 500 Index) to be null. Demeterfi et al. in [13] (1999) showed an initial review of the Eq. (3.62). They provided an alternative definition for VIX that places it in direct correlation with the volatility of the model replicating the SPX. Considering as a generic model

$$\frac{dS_t}{S_t} = rdt + \sigma(S_t)S_t dW_t \tag{3.63}$$

for the SPX, we can define the VIX as only a function of $\sigma_t (:= \sigma(S_t))$[3]:

$$VIX_t^2 = E_t^Q\left[\frac{1}{\tau}\int_t^{t+\tau} \sigma_s^2 ds\right] \tag{3.64}$$

[3]Use of risk-neutral probability measure (Q-measure).

where τ indicates the 30-day time interval typical of the VIX. In practice, this formula does not directly link the VIX index to the value of the underlying asset (SPX). Yingzi Zhu in "Variance Term Structure and VIX Futures Pricing" ([40]) developed the steps to rewrite Eq. (3.63) by highlighting the link between the two indices. We now apply Itô's lemma:

$$d\ln S_t = rdt - \frac{1}{2}\sigma_t^2 dt + \sigma_t dW_t \tag{3.65}$$

that combined with (3.63) becomes

$$d\ln S_t = \frac{dS_t}{S_t} - \frac{1}{2}\sigma_t^2 dt \tag{3.66}$$

If we now integrate over a time interval $[t, t+\tau]$ we gain:

$$\int_t^{t+\tau} d\ln S_t = \ln S_{t+\tau} - \ln S_t = \int_t^{t+\tau} \frac{dS_t}{S_t} - \frac{1}{2}\int_t^{t+\tau} \sigma_t^2 dt \tag{3.67}$$

which we can rewrite as

$$\int_t^{t+\tau} \sigma_t^2 dt = 2\int_t^{t+\tau} \frac{dS_t}{S_t} - 2\left(\ln S_{t+\tau} - \ln S_t\right) \tag{3.68}$$

Let's focus on the stochastic integral $\int_t^{t+\tau} \frac{dS_t}{S_t}$: we can use in fact the definition given in (3.63) and develop the integral as:

$$\int_t^{t+\tau} \frac{dS_t}{S_t} = \int_t^{t+\tau} rdt + \int_t^{t+\tau} \sigma_t dW_t \tag{3.69}$$

Considering a zero interest rate as in our case, the integral turns out to be zero; for the second integral, namely the stochastic one, the reasoning is more complex. The integral alone, having a Brownian motion as driver turns out to be a martingale, and thus does not allow us to obtain an effective result. Knowing, however, that the starting integral is contained in an expected value and that the linearity rule applies, we can rewrite (3.69) as:

$$\begin{aligned} E_t\left[\int_t^{t+\tau} \frac{dS_t}{S_t}\right] &= E_t\left[\int_t^{t+\tau} rdt\right] + E_t\left[\int_t^{t+\tau} \sigma_t dW_t\right] \\ &= E_t\left[\int_t^{t+\tau} \sigma_t dW_t\right] \end{aligned} \tag{3.70}$$

Now, since we have the expected value of a martingale[4], we conclude that its value is zero. So (3.68) becomes

$$\begin{aligned} E_t\left[\int_t^{t+\tau}\sigma_t^2 dt\right] &= 2E_t\left[\int_t^{t+\tau}\frac{dS_t}{S_t}\right] - 2E_t\left[(\ln S_{t+\tau} - \ln S_t)\right] \\ &= -2E_t\left[(\ln S_{t+\tau} - \ln S_t)\right] \\ &= 2E_t\left[\ln(S_t)\right] - 2E_t\left[(\ln S_{t+\tau})\right] \end{aligned} \tag{3.71}$$

recalling that the log-price by definition is X, we can redefine the VIX on the basis of X, thus going on to emphasize the direct link between the volatility index and the asset price on which it is based:

$$\begin{aligned} VIX_t &= \sqrt{E_t\left[\frac{1}{\tau}\int_t^{t+\tau}\sigma_s^2 ds\right]} \\ &= \sqrt{\frac{2}{\tau}\left(E_t\left[\ln(S_t)\right] - E_t\left[(\ln S_{t+\tau})\right]\right)} \\ &= \sqrt{\frac{2}{\tau}\left(E_t\left[X_t\right] - E_t\left[X_{t+\tau}\right]\right)} \\ &= \sqrt{\frac{2}{\tau}\left(X_t - E_t\left[X_{t+\tau}\right]\right)} \end{aligned} \tag{3.72}$$

So to simulate the VIX at a given instant of time t, one must simulate the log-price of the underlying asset up to t and then from that moment again up to time $t+\tau$.

We begin by considering the case in which the underlying is simulated using the Monte Carlo method. Let us say that we want to simulate only one path of VIX_t up to time t: first we produce a single simulation of the underlying (the log-price) up to this instant exactly as exposed in the previous sections, producing a value for X_t. To compose the remaining part of the index we need to compute $E_t\left[X_{t+\tau}\right]$, which is the expected value of the simulations of the log-price process starting t, up to $t+\tau$. At the practical level, to do this, the information at time t of all the dynamics that make up the model is taken and used as initial conditions for the secondary simulations of the process; having obtained a set of dynamics, with the expected value we calculate only one value of VIX_t. This process must be simulated multiple times in order to be used in reality to ensure that the Monte Carlo method is sufficiently stable and accurate. It can therefore be seen immediately how time-consuming this method is: for each simulation of the log-price up to time t, additional simulations must be performed to complete the calculation of

[4]Recall that a martingale is called a stochastic process $(Y_t)_{0\le t\le T}$ such that each Y_t is $\mathcal{F}_t$-measurable and integrable and that, taken $s<t$, the equality $Y_s = E\left[Y_t \mid \mathcal{F}_s\right]$ holds.

VIX_t. This algorithm is called "*nested loop*" and, calling n_1 the number of simulations of the initial loop, and n_2 the number of simulations of each inner loop, it needs a total of $n_1 \cdot n_2$ paths in order to compute the vector of different paths for the VIX_t. Obviously, in order to obtain accurate results, it turns out that the total number of simulations to be performed must be very high, and consequently high computational times are expected.

Considering instead the binomial tree as a pricing technique, the matter is different. The concept of simulating X_t and later $X_{t+\tau}$ starting from t persists, but the method changes completely. In particular, thanks to the non-recombinant tree structure, by simulating the log-price of the model up to time $t+\tau$ it is possible to obtain all the information we need. In fact, thanks to the particular structure of this model, one can see the simulation tree of $X_{t+\tau}$ as an initial tree going from 0 to t to which, for each node obtained, an additional tree with such starting conditions is created. Conceptually this seems analogous to simulation by the Monte Carlo method, and indeed it is, but on a practical level it is not. In fact considering a generic simulation X_t, contained in that of $X_{t+\tau}$, is already present the set of paths to get to $t+\tau$ starting from t to compute $E_t[X_{t+\tau}]$. Considering all nodes present at instant t of the single generated tree, it is then possible to create a vector of simulations for VIX_t to be used later in calculating derivative prices. Thanks to the recombinant tree, by developing two simulations starting from the same instant t, it is not possible for the paths to be in common. This makes it easy to identify the development of a model dynamics from a given instant onward. If the tree had been recombinant, however, highlighting the growth of a dynamic (the secondary tree) would not have been trivial. In fact, more attention would have had to be paid to actually tracking the development of a given node, even if the dynamics intersect with those of the paths of other nodes.

Both methods then simulate the value of the log-price of the model up to time t and then for each result obtained, a further series of simulations is carried out up to time $t+\tau$. The abysmal practical difference between the two techniques lies in the fact that for Monte Carlo it is necessary to apply the "*nasted loop*", while for the recombinant binomial tree method, it is sufficient to directly simulate the log-price dynamics in the interval 0, $t+\tau$.

4 | Analysis of pricing results

Having described the model from a theoretical point of view and analyzed it from an application perspective, in this chapter we move on to analyze the actual results of applying the model. In particular, we will go on to compare the various techniques described, highlighting strengths and weaknesses of the pricing techniques that can be used. We will mainly refer to the calculation of European call and put options through which implied volatility can be calculated: it's fundamental since, in real application in the financial world, it is very important to be able to describe the smile of implied volatility.

4.1. General dynamics of the model

We now want to take a general look at what the results obtained are with respect to the model's ability to reproduce a financial performance. Specifically, by going to replicate the value of the log-price, volatility, and R_1 and R_2 features, we want to emphasize the range of existence of certain values.

4.1.1. Dynamics with Monte Carlo

We begin by analyzing the dynamics produced by the Monte Carlo method, applying variance reduction technique "antithetic variable" and using Milstein discretization since σ_t has non-trivial dynamics. Figure 4.1 shows the trends of the four dynamics of the 4-factor model (actually, instead of $R_{1,0,t}$, $R_{1,1,t}$, $R_{2,0,t}$, $R_{2,1,t}$, are used $R_{1,t}$ and $R_{2,t}$), starting from the values introduced by Guyon and Lekeufack in [25] shown in Table 4.1.

Parameters value of the 4-factor model

$\lambda_{1,0}$	$\lambda_{1,1}$	$\lambda_{2,0}$	$\lambda_{2,1}$	β_0	β_1	β_2	θ_1	θ_2
55.0	10.0	20.0	3.0	0.04	-0.13	0.65	0.25	0.5

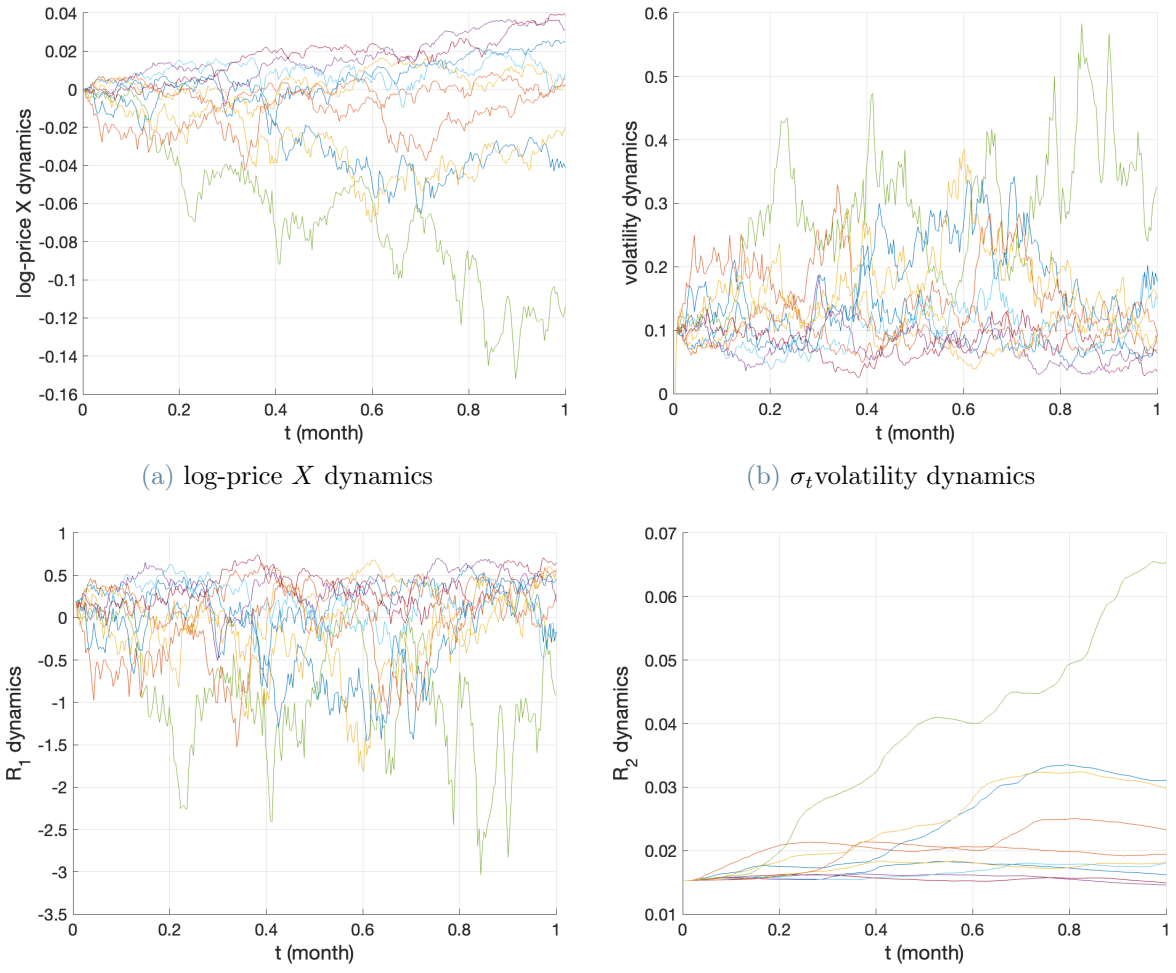
Table 4.1: Set of parameters selected by Guyon and Lekeufack for the simulation of the 4-factor Markovian PDV model

With respect to the initial values of the variables with dynamics instead, we used the values presented in Table 4.2.

Initial value of the 4-factor model variables

X_0	$R_{1,0,0}$	$R_{1,1,0}$	$R_{2,0,0}$	$R_{2,1,0}$
0.0	0.168	0.244	0.005	0.003

Table 4.2: Initial values of dynamic variables in the 4-factor Markovian PDV model



(a) log-price X dynamics

(b) $\sigma_t$volatility dynamics

(c) R_1 dynamics

(d) R_2 dynamics

Figure 4.1: Dynamics of 10 simulations of the 4-factor model developed over a period of 1 month by Monte Carlo method with initial value from Table 4.1 and Table 4.2. We used 10 steps per day.

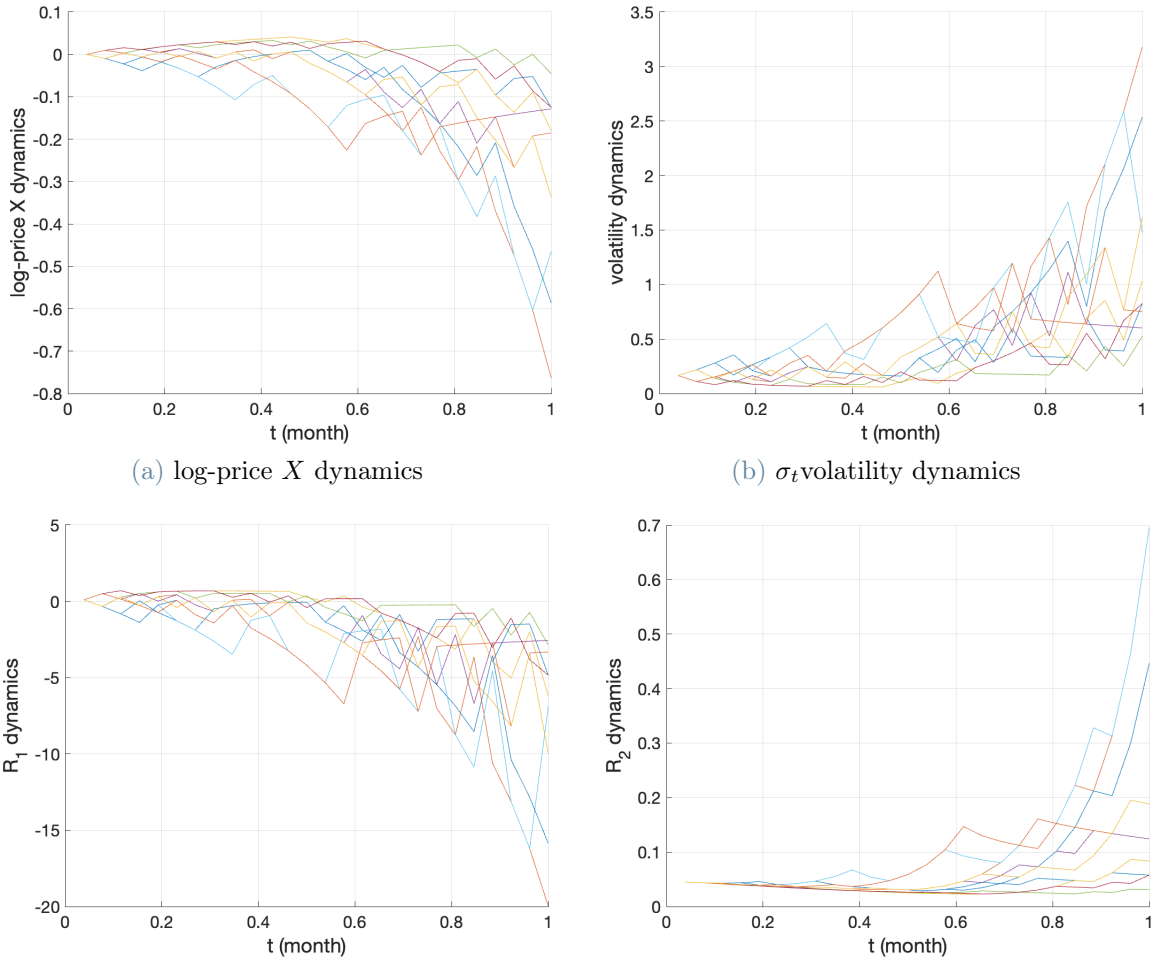
Looking at Figure 4.1, it is possible to deduce and confirm concepts already expressed

and explored in Chapter 2. Looking at the X-dynamics, it can be seen that most of the routes after one month reach similar values to each other, and on average slightly higher than the starting value. Only a few dynamics differ from this trend: thanks to these, it is possible to see how volatility values, in correspondence to these events, tend to increase. The dynamics of R_1, as anticipated earlier, is able to objectively explain the Leverage Effect, (the fact that volatility, in the presence of a decline in asset value, tends to increase), giving a very positive contribution, through the negative coefficient β_1, to the value of volatility. The capacity, on the other hand, for volatility clustering, that is the tendency to follow periods characterized by high volatility with equally high volatility and periods of low volatility with equally low volatility, is provided by the presence of R_2. In fact, this dynamic, compared with R_1, is much less volatile with an existence range generally between $[0, 2]$, quite different from R_1, which is generally in the range $[-15, 4]$ instead. Even if obvious, it is important to note that the values of R_2, which make up volatility through a square root function, turn out to be always positive and therefore volatility also turns out to be so. The purely negative R_1 contributions are balanced by a negative β_1 coefficient.

4.1.2. Dynamics with Binomial Tree

We have seen the results produced by the Monte Carlo method, let us continue with those produced by the binomial tree method. Let us focus on the differences involved in adopting the binomial approach: it is possible to see immediately in Figure 4.2, despite the fact that 10 different paths were randomly chosen from each other, a lattice structure within the evolution of the dynamics. This effect obviously tends to be more evident by increasing the number of simulations, at the expense of large computational costs. This structure is obviously due to the possibility, at a given instant of time, of being able to continue only along two paths, concept behind the basic of the binomial tree. Despite this trend, however, the considerations made with respect to all dynamics are confirmed: R_1 remains mostly negative, R_2 and σ_t take only positive values, and the log-price struggles to grow, preferring paths along which it decreases in value.

Since the binomial tree is a non-recombinant tree in this case, at each time step for each value found the previous instant, two separate values will be generated for each node that are not linked together. This makes each vector of dynamics equal in length to 2^n, where n denotes the time step at which we are. It is clear that despite the fact that the number of steps may not be too large, it is possible to obtain high size vectors. Obviously from the point of view of computational cost this is not insignificant since high time discretization cannot be achieved. Thus, the binomial model, like other techniques used in the pricing



(a) log-price X dynamics

(b) $\sigma_t$volatility dynamics

(c) R_1 dynamics

(d) R_2 dynamics

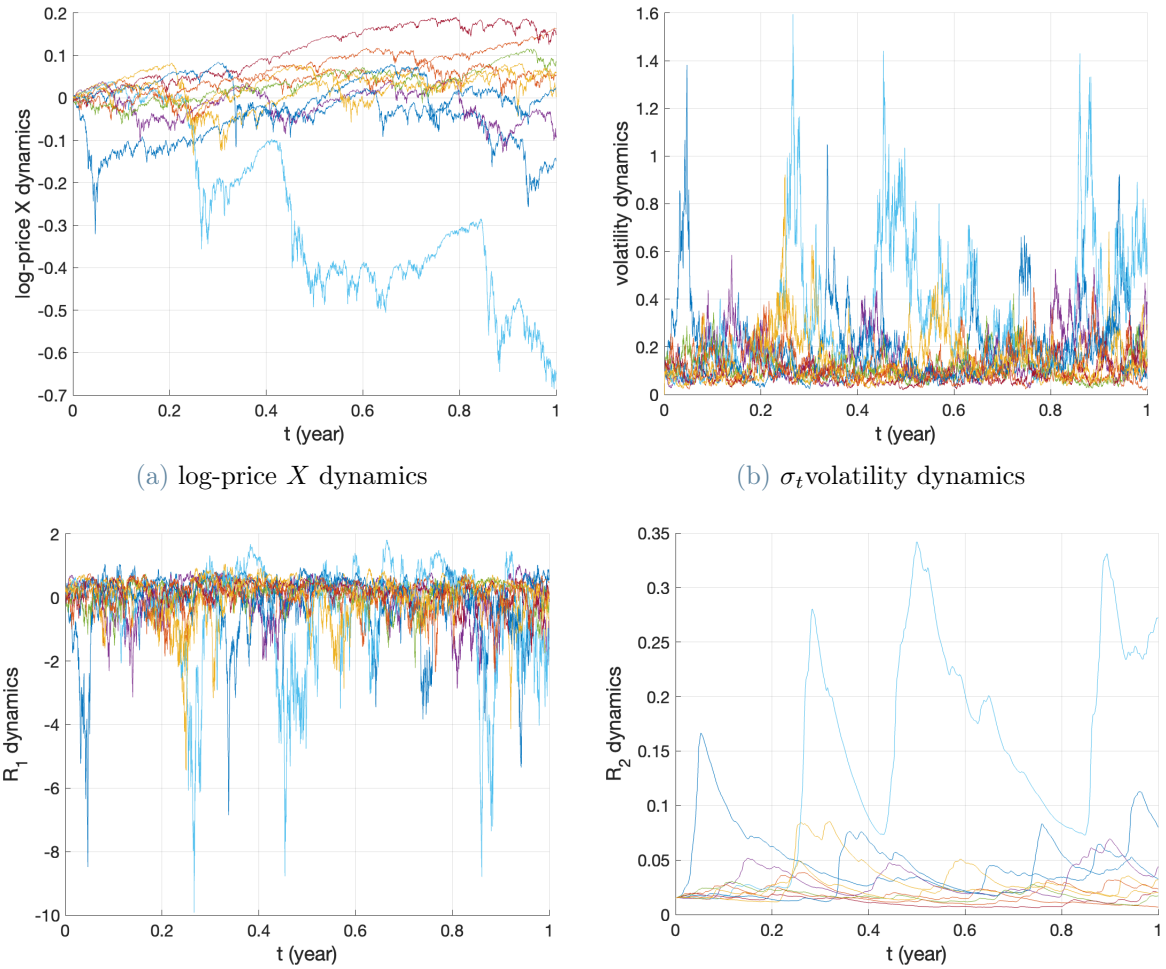
Figure 4.2: Dynamics of 10 simulations of the 4-factor model developed over a period of 1 month by binomial tree method with initial value from Table 4.1 and Table 4.2. We used 26 total steps.

world, appears to be ineffective in representing the performance of the underlying despite the fact that, from a pricing perspective for derivatives, it can be a viable alternative. For completeness of analysis, we then go on to observe over a one year time frame the behavior of the model dynamics; we therefore use of the Monte Carlo method.

4.1.3. 1-year Monte Carlo dynamics

In the previous section, we analyzed the dynamics of the model and their significance over a time frame of one month in order to compare the results of the Monte Carlo method and the binomial tree. As explained the binomial tree has a limitation with respect to the time frame in which it can be adopted. We present below in Figure 4.3 the dynamics

of the 4-factor model for a time frame of one year. As can be seen, the considerations made in Section 4.1.1 and 4.1.2 are confirmed.



(a) log-price X dynamics

(b) $\sigma_t$volatility dynamics

(c) R_1 dynamics

(d) R_2 dynamics

Figure 4.3: Dynamics of 10 simulations of the 4-factor model developed over a period of 1 year by Monte Carlo method with initial value from Table 4.1 and Table 4.2. We used 10 steps per day.

All results with respect to the dynamics we have seen come from the study of the 4-factor model, as it is more complete than the 2-factor model where the TSPL kernel is not well represented. However, the 2-factor model, as Figure 2.3 shows, despite having a simple exponential kernel, manages to replicate the TSPL kernel well for short times. It therefore turns out to be an efficient model when it comes to pricing for derivatives with short maturities.

4.2. Pricing results

In this section we want to present, compare, and graphically represent the results of the smiles of implied volatility of both the SPX and the VIX. To do this we will use European options: we will need to calculate the price of European calls, and consequently the implied volatility, for strikes above the price of the underlying asset (out of the money, OTM), and European puts for strikes below the price of the stock (in the money, ITM). We are going to consider different strikes for different maturities: strikes will vary from 75% to 104% of the underlying for options with a maturity of one month, to values from 75% to 120% for options with a maturity of one year. The same strikes that Guyon and Lekeufack used in their paper were chosen. To calculate the implied volatility, having obtained the price of the derivative and knowing the underlying R_0, the interest rate r, and the maturity T, it is possible to use Matlab[1].

We have already discussed the limitation of time steps in the binomial tree method. However, there is a solution to expand this limitation: the nested binomial tree approach. In this case, each node created at a specific time triggers the generation of a new tree using the conditions present at that time as a starting point. As a result, this method allows a significant increase in the number of time steps. Although not used in the results shown in Figure 4.2, this technique is applicable to European options and remains remarkably fast and efficient.

4.2.1. 2-factor model: pricing result

Let us introduce the value of the parameters to be adopted in the 2-factor model presented by Guyon and Lekeufack in their paper [25] in Table 4.3.

Parameters value of the 2-factor model

λ_1	λ_2	β_0	β_1	β_2
62.0	40	0.08	-0.08	0.5

Table 4.3: Set of parameters selected by Guyon and Lekeufack for the simulation of the 2-factor Markovian PDV model

The initial value of the 2-factor model are summarized in Table 4.4.

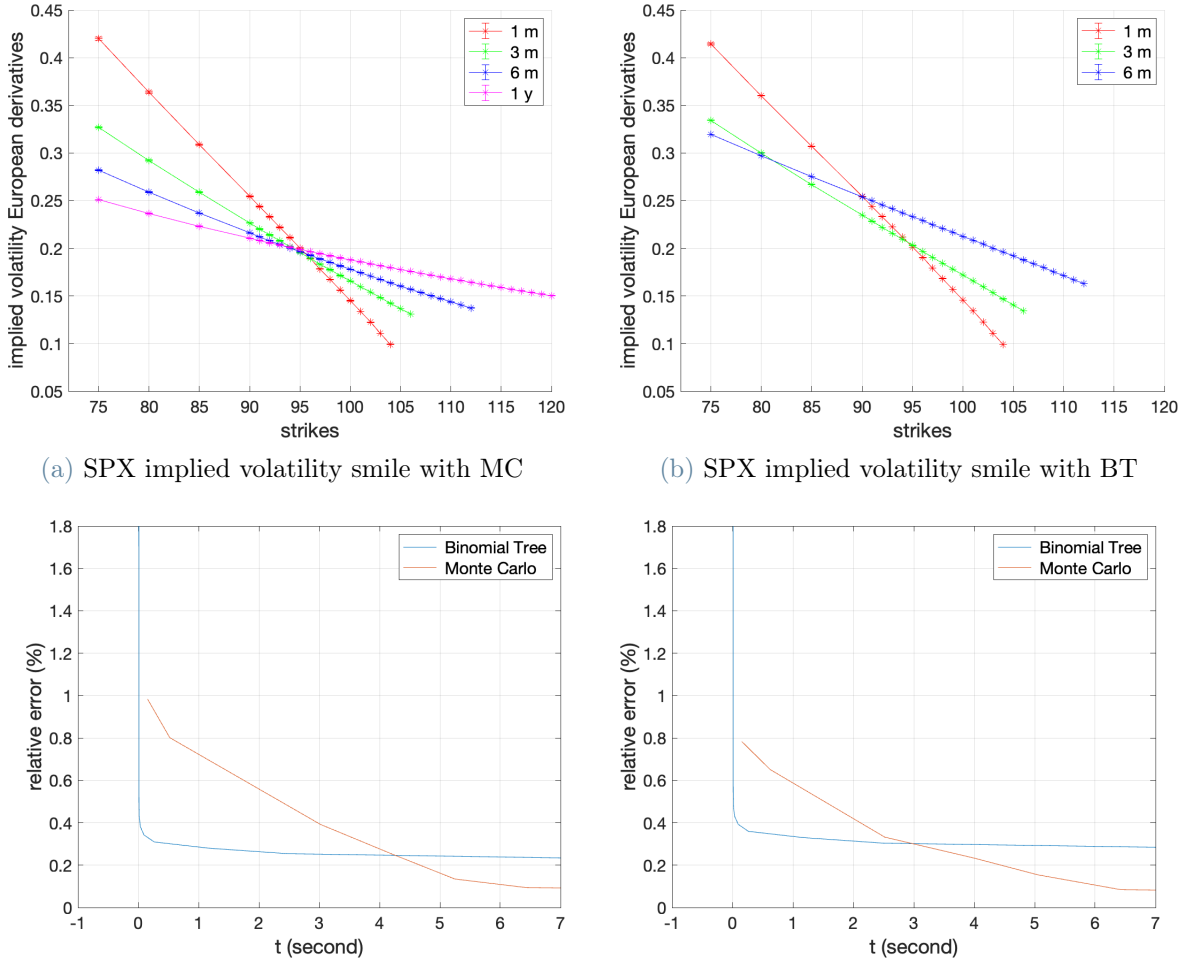
[1]In particular is used *blkimpv*: function to compute the Implied volatility for futures options from Black model on the Matlab calculator.

Initial value of the 2-factor model variables

S_0	$R_{1,0}$	$R_{2,0}$
100.0	-0.044	0.007

Table 4.4: Initial values of dynamic variables in the 2-factor Markovian PDV model

We present the results graphically in Figure 4.4:



(a) SPX implied volatility smile with MC

(b) SPX implied volatility smile with BT

(c) Relative error and computational time

(d) Relative error and computational time

Figure 4.4: General results related to the SPX implied volatility results for the 2-factor model. Values and comparison in computational times of Monte Carlo and binomial tree methods. Initial value for the model from Table 4.3 and Table 4.4.

We start by analyzing Figure 4.4a in which are shown the implied volatility values for European derivatives calculated from an underlying simulated by me Monte Carlo Carlo

method. Specifically, to obtain the results presented, 10 different simulations were run where each was composed of 10^8 paths and 10 time steps per day; an average was then performed to obtain the derivative prices and then the implied volatility smile for SPX. This process was performed once for each maturity analyzed (1 month, 3 months, 6 months, and 1 year), and given the large number of paths to obtain a minimum error, the time for each is more than 10 minutes. Derived prices obtained by this method were chosen as benchmarks for the 2-factor model for future comparisons.

In Figure 4.4b we always have the implied volatility smile for the 2-factor model, but calculated via the binomial tree. As previously suggested the main limitation for this method concerns the number of total steps that can be consider; in fact, since the tree is non-recombinant, the length of the vectors used for the simulation tends to grow exponentially. Through nested paths, however, it is possible to reach 38 total time steps (running only one nested loop) and still obtain good results in reasonable time. This technique, on the other hand, is successful only for relatively short maturities; as can be seen for maturities of one month and three months the values obtained are consistent with those in Figure 4.4a, for but six months or more (not shown) the difference between the two methods is obvious.

The financial world, and more particularly the world of trading, is becoming faster and more dynamic. This phenomenon makes pricing results obtained in long time frames of low interest. We therefore analyze errors and computational times of the two models using the results in Figure 4.4a as benchmarks (we compare only data with maturities of 1 month and 3 months). To compare the two models we will use the relative error (in percentage) with respect to the benchmark respectively as the number of simulation and daily steps increases for the Monte Carlo method and the number of total steps for the binomial tree method. For the Monte Carlo method, there will be no multiple simulations used and then averaged but directly one.

Referring to Figure 4.4c, we observe how the error varies over time with respect to the price of the derivative with maturity one month. Considering that the error calculated by the binomial method is inversely proportional to the number of intervals ($\frac{1}{n}$) and that the computational time is linear with the number of steps, it appears as in the figure that the error is inversely proportional to time ($\frac{1}{t}$). For the Monte Carlo method, on the other hand, the error is inversely proportional to the square root of the number of simulations applied ($\frac{1}{\sqrt{N_{sim}}}$), and in this case, considering the relationship with time, we obtain a trend, even if more rough, inversely proportional to the root of time ($\frac{1}{\sqrt{t}}$). For computational times below approximately four seconds, the binomial tree method brings better results per decimal percentage point, which, with large investments, can be crucial.

For longer computational times, on the other hand, the Monte Carlo method is better and succeeds in getting closer to zero. The area between the two curves at the initial instants is an indication of how much better one method forms than the other. In fact by increasing maturity to three months (Figure 4.4d) this area decreases. In this case, the best discretization that can be applied by the binomial tree method leads to an overall increase in relative error. This causes that the point of convenience between one method and the other is at three seconds, not four as the previous case.

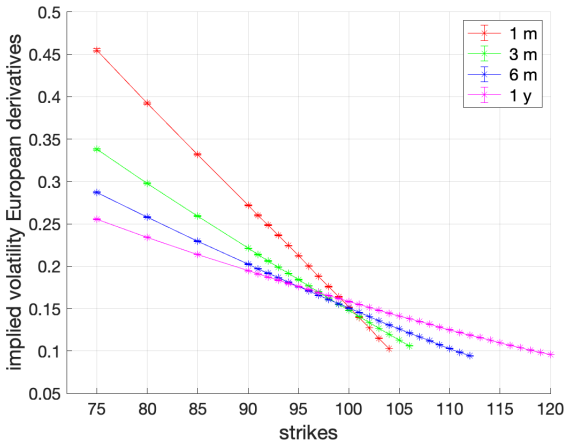
4.2.2. 4-factor model: pricing result

We now proceed with the analysis and comparison of the results obtained through the 4-factor model, which will be similar to the 2-factor model since for short time intervals the kernels of both models approximate very well the TSPL kernels. As before in Figure 4.5 we show the results. We use initial value from Table 4.1 and Table 4.2. Figure 4.5a and 4.5b show very similar values for maturities of 1 month and 3 months; from 6 months on, however, the results diverge very clearly.

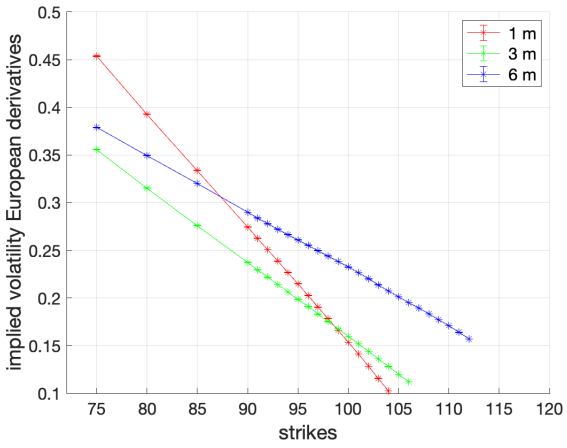
By analyzing Figure 4.5c and 4.5d, instead, it is possible to expand on the previous considerations. Comparing the computational times with respect to the 2-factor method, for both maturities calculated by binomial tree method, it is possible to observe that the relative error is greater given an instant of time. This is mainly due to the fact that the number of dynamics on which another non-recombinant tree must be constructed increases. This phenomenon is less evident in the application of the Monte Carlo method since, although the number of dynamics increases, the simulations vector of the stochastic component of the PDV model does not change. This causes that the area between the two error curves in the 4-factor model is smaller than in the 2-factor model. It also decreases more by going from Figure 4.5c and Figure 4.5d; this is confirmed by the fact that the convenience points between one method and the other drop respectively to 3 and 2.4 seconds. The trends of relative error in relation with time confirm the arguments made earlier.

It should be noted that for both methods this model results in a significant increase in computational cost in terms of memory. Simulations obtained with Monte Carlo are affected by a simple increase in the number of variables. The non-recombinant tree, on the other hand, involves the addition of two new integer diagrams: considering that the vectors can grow to sizes of more than 100 million values, the memory of the calculator is put to a serious effort. For this reason, again considering short deadlines, the 2-factor model may be a viable alternative because, having fewer dynamics, it requires less computational

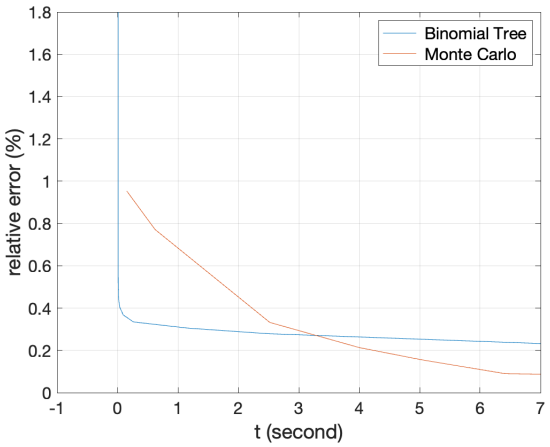
cost.



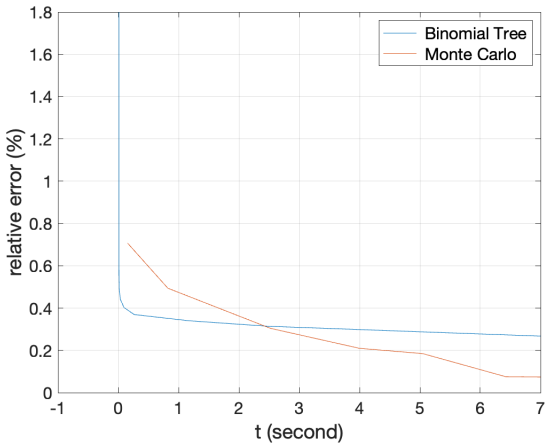
(a) SPX implied volatility smile with MC



(b) SPX implied volatility smile with BT



(c) Relative error and computational time



(d) Relative error and computational time

Figure 4.5: General results related to the SPX implied volatility results for the 4-factor model. Values and comparison in computational times of Monte Carlo and binomial tree methods. Initial value for the model from Table 4.1 and Table 4.2.

4.2.3. General considerations

Analyzing the implied volatility smile for the SPX, it can be affirmed that the binomial tree method turns out to be a valid alternative to Monte Carlo simulations, which are very popular in the financial world. For short maturities in fact it produces the best results in less time. Despite the fact that these are seconds, in today's rapid financial transactions, they can make the difference. The best computational time, however, comes at the expense of computational cost. In order to apply this method to both the 2-factor and 4-factor models, it is necessary to allocate a large amount of memory for the variables.

4.3. The big step: VIX implied volatility smile

One of the main achievements introduced by the PDV model is the ability to simulate the smiley of the implied volatility of the VIX in a manner very similar to empirical data. This problem, is addressed by using nested loops and Monte Carlo simulations (the approach proposed in Guyon and Lekeufack (2023) [21] to price VIX derivatives); in this the results are accurate, but with a high computational cost in terms of memory and computation time. In order to analyze the values graphically, we simulate the VIX index at one month and two months (maturities on which the binomial tree model is found to work) with the 2-factor model. We use initial value from Table 4.3 and Table 4.4. Knowing that the value of VIX is about 0.2 and having produced the simulations up to maturity, we calculate calls and puts with strikes ranging from 0.15 to 0.5 . As in the previous section, to compute the smile of implied volatility, we use "out of the money" European calls and "in the money" European puts.

4.3.1. VIX derivatives

In Section 3.4.1 we highlighted the direct link between the value of the VIX and the price of the underlying asset to which it refers. The final goal is to understand what results the model introduced in this Thesis brings with respect to the smile curve of implied volatility for the variance index. First we need to calculate the price of calls and puts for different strikes given an expiration date. We then proceed similarly for a generic underlying, specifically for European call options we will have

$$C(\text{VIX})_{EU} = E_0\left[(VIX_T - K)^+\right] \tag{4.1}$$

where T denotes the maturity and K the strike. Similarly, we deduce the formula for the European put:

$$P(\text{VIX})_{EU} = E_0\left[(K - VIX_T)^+\right] \tag{4.2}$$

For both derivatives, it is therefore necessary to simulate only the VIX up to time T. Then, through a simple expected value that takes into account the discount factor (null in our case since r is zero), obtain the option prices. Recall that given a certain maturity, to simulate the VIX up to that time (T), the log-price X_t must be simulated up to instant $T + \tau$ for the tree model, while for the Monte Carlo Method, starting from T, we have to resimulate to arrive at $T + \tau$.

4.3.2. VIX results

The nested loops used for the Monte Carlo method occupy a very large amount of memory and do not allow the use of sufficiently large vectors for simulations. This results in a very high error and especially much related to the random component to be attributed to the stochastic variables in the model. Nested cycles, because of their structure, take a long time to produce results; the implied volatility curve takes times longer than 7 minutes for each expiration.

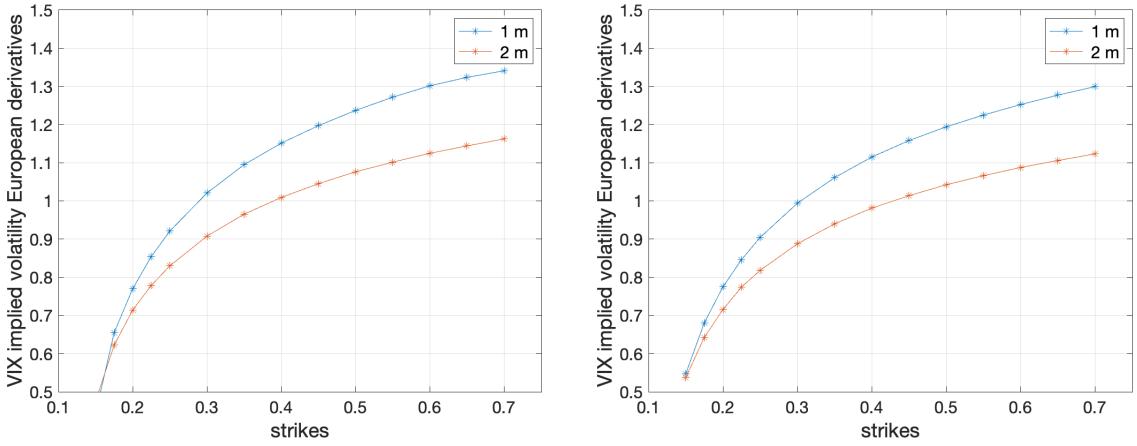
We now discuss the results obtained from the binomial tree (Figure 4.6b), a new method for dealing with the pricing of Guyon and Lekeufack's model. Through the tree structure in fact, given a certain instant of time, a smaller tree of equal size is created for each value. This peculiarity is very convenient in the simulation of the VIX since to calculate its value in one month's time, we replicate for two months the value of SPX and then, at the midpoint of the time interval, considers the beginning of the 30 days typical of the VIX. The same logic can be applied for the two-month VIX: we simulate the SPX for three months and consider the values at 2 and 3 months. Thus, having in the binomial tree already available the SPX values needed for the mix calculation, the computational times are very low: only 11 seconds per maturity (see Table 4.5).

Computational timing for the VIX implied volatility smile.

	1-month maturity	**2-months maturity**
Monte Carlo	6.8 min	7.1 min
Binomial Tree	8.54 sec	10.96 sec

Table 4.5: Computational timing for different methods and maturities

The values presented in Figure 4.6 are different from each other because both methods have weaknesses related to the calculator that is used. In particular, considering the simulations generated with the Monte Carlo method, we obtain results with very high variance: by increasing the number of simulations of the first loop and the nested loop, we can get more accurate results at the expense of time and memory. Compared with the use of the binomial tree, a low number of time steps used, despite the fact that short deadlines are involved, can be a limitation: with more memory, it is possible to increase temporal discretization without possibly exceeding computational time. However, for strikes above the value of the VIX and relatively close to the ATM point ("at the money"), the two methods present very similar results.



(a) VIX implied volatility smile with MC (b) VIX implied volatility smile with BT

Figure 4.6: General results related to the VIX implied volatility results for the 2-factor model, considering 1 month and 3 months as maturities. Use of 20,000 paths and 5,000 nested paths for Monte Carlo and 38 total time steps (running only one nested loop) for binomial tree method. Initial value for the model from Table 4.3 and Table 4.4.

5 | Conclusions and future developments

Let us draw conclusions with respect to the model and methods used.

5.1. Conclusions

The purpose of this Thesis is to understand if the PDV model of Guyon and Lekeufack (2022) can be used to price efficiently derivative instruments. Three are the main contribution of this Thesis.

First, we explain in detail how to implement a Monte Carlo scheme to price options on the S&P 500 and VIX in the framework of the PDV model. We improve the existing method described by Guyon and Lekeufack (2022) by proving how adapt the Milstein scheme (see Section 3.3.2) to the case of path dependent volatility models. In particular, the PDV model has been designed to solve the open problem of a joint calibration of S&P 500 and VIX volatility surface. Computing prices of option on the underlying with Monte Carlo simulations is well established on the literature. Computing the price of options on the VIX is less so. We explain which is the relevant quantity to simulate when dealing with VIX options (see Section 3.4) and why it is necessary recurring to an inefficient nested Monte Carlo simulation.

Second, we propose an alternative method for pricing options on the S&P 500 and VIX: a non-recombinant binomial tree. The structure of the proposed trees for the 2 and 4-factor PDV models is described in Section 3.3.4. We start by deducting the recombinant binomial tree for Brownian Motion only, a process with simple dynamics. We then develop the tree for each dynamics of the 2 and 4-factor models. We thus obtain non-recombinant binomial trees, which require some attention in practical development. Although more complex to implement, they bring an advantage in the calculation of VIX. Indeed, the lattice structure with non-overlapping paths is optimal for the simulations required to determine the value of VIX without additional modifications.

Finally, we compare the performance of the two methods. We compare the error obtained with the two methods. The binomial tree method, when dealing with S&P 500 derivatives, succeeds in obtaining better results for times less than about 4 seconds (for option with maturity below six months), with an error around 20 basis points. For errors around the basis point (industry standard) the Monte Carlo remains more efficient, even if the computational time increases. When we want to compute an option on the VIX, however, the computational times of the tree is far less time-consuming than those found with the Monte Carlo method (on average 11 seconds instead of 7 minutes) for the Monte Carlo method. Finally, we point out how when increasing the maturity of options on S&P 500 or VIX over six months both methods become highly inefficient at least with standard computational power.

5.2. Future developments

Let us now look at some solutions and/or ideas can be applied to investigate the model more in order to obtain more accurate results. Below are some suggestions:

- **Larger memory**: although it may seem trivial to have a larger memory capacity, it can significantly increase the accuracy of the results. In the Monte Carlo method, this increase would allow for greater discretization of time and especially the possibility of having simulations composed of multiple values. Particularly in the calculation of the smile of the VIX implied volatility, it would be possible to achieve the minimum number to ensure the stability of the result. Referring instead to the method with the binomial tree, more memory would allow a larger number of time steps, since at each of them the length of the vector of simulations of the dynamics of the process is of size 2^n. This increase would already be evident in the calculation of the derivatives for the S&P 500[1].
- **Clustering of the lattice structure**: if an increase in memory is not possible for the binomial method, a technique for increasing time steps can be used. Clustering at each instant the values of the simulations, and it is possible to obtain vectors of reduced size. In particular, an analysis of the dynamics of the process for each time instant would be performed, agglomerating data of very similar value. Obviously, by doing so, the tree would no longer be non-recombinant, so, especially for the calculation of VIX, more attention would have to be paid to the dynamics arising from a specific node and not to confuse them with those of other neighboring nodes.

[1]For the results shown in this Thesis, a computer with a 2.6 GHz Intel Core i7 6-core processor and 16 GB 2400 MHz DDR4 memory has been used.

- **Neural networks**: neural networks constitute a significant evolution in the algorithmic scene, bringing with them significant advances in time efficiency and computational cost. These first results were shown by Gencay and Qi in [19] and later by Gradojevic, Gencay and Kukolj in [23]. These improvements arise mainly from their ability to manipulate complex representations of data. The use of neural networks enables massive parallelization of operations, taking full advantage of modern hardware architectures, such as GPUs. This parallelization speeds up the calculating process, allowing larger amounts of data to be handled in considerably less time. This translates into the ability to build more sophisticated models without exponentially inflating computational costs; transposing this concept to the tree structure theoretically opens up possibilities for improving the accuracy of results.

Bibliography

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

[2] V. Bally and D. Talay. The law of the euler scheme for stochastic differential equations: I. convergence rate of the distribution function. *Probability Theory and Related fields*, 104:43–60, 1996.

[3] O. E. Barndorff-Nielsen. Processes of normal inverse gaussian type. *Finance and Stochastics*, 2:41–68, 1997.

[4] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of political economy*, 81(3):637–654, 1973.

[5] P. Blanc, J. Donier, and J.-P. Bouchaud. Quadratic hawkes processes for financial prices. *Quantitative Finance*, 17(2):171–188, 2017.

[6] J.-P. Bouchaud, A. Matacz, and M. Potters. Leverage effect in financial markets: The retarded volatility model. *Physical Review Letters*, 87(22):228701, 2001.

[7] G. Brunick and S. Shreve. Mimicking an itô process by a solution of a stochastic differential equation. 2013.

[8] M. Chesney and L. Scott. Pricing european currency options: A comparison of the modified black-scholes model and a random variance model. *Journal of Financial and Quantitative Analysis*, 24(3):267–284, 1989.

[9] F. Comte and E. Renault. Long memory in continuous-time stochastic volatility models. *Mathematical finance*, 8(4):291–323, 1998.

[10] J. C. Cox, S. A. Ross, and M. Rubinstein. Option pricing: A simplified approach. *Journal of Financial Economics*, 7(3):229–263, 1979.

[11] J. C. Cox, J. E. Ingersoll Jr, and S. A. Ross. An intertemporal general equilibrium model of asset prices. *Econometrica: Journal of the Econometric Society*, pages 363–384, 1985.

[12] P. Del Moral and P. Del Moral. *Feynman-kac formulae*. Springer, 2004.

[13] K. Demeterfi, E. Derman, M. Kamal, and J. Zou. More than you ever wanted to know about volatility swaps. *Goldman Sachs quantitative strategies research notes*, 41:Appendix B, 1999.

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

[15] B. Dupire et al. Pricing with a smile. *Risk*, 7(1):18–20, 1994.

[16] C.-C. B. O. Exchange. Cboe VIX: White Paper Cboe Volatility Index, 2019.

[17] C.-C. B. O. Exchange. Cboe VIX: Volatility Index Methodology: Cboe Volatility Index, 2022.

[18] M. Fukasawa, T. Takabatake, and R. Westphal. Is volatility rough? *arXiv preprint arXiv:1905.04852*, 2019.

[19] R. Gençay and M. Qi. Pricing and hedging derivative securities with neural networks: Bayesian regularization, early stopping, and bagging. *IEEE transactions on neural networks*, 12(4):726–734, 2001.

[20] D. T. Gillespie. Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral. *Physical review E*, 54(2):2084, 1996.

[21] P. Glasserman. *Monte Carlo methods in financial engineering*, volume 53. Springer, 2004.

[22] P. Glasserman, P. Heidelberger, and P. Shahabuddin. Variance reduction techniques for estimating value-at-risk. *Management Science*, 46(10):1349–1364, 2000.

[23] N. Gradojevic, R. Gençay, and D. Kukolj. Option pricing with modular neural networks. *IEEE transactions on neural networks*, 20(4):626–637, 2009.

[24] J. Guyon. Path-dependent volatility. *September 10, 2014, Available at SSRN: https://ssrn.com/abstract=2425048*.

[25] J. Guyon and J. Lekeufack. Volatility is (mostly) path-dependent. *Volatility Is (Mostly) Path-Dependent (July 27, 2022)*, 2022.

[26] S. L. Heston. 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, 1993.

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

[28] P. Jäckel. *Monte Carlo methods in finance*, volume 5. John Wiley & Sons, 2002.

[29] R. Jarrow and A. Rudd. Approximate option valuation for arbitrary stochastic processes. *Journal of Financial Economics*, 10(3):347–369, 1982.

[30] C. Kahl and P. Jäckel. Fast strong approximation monte carlo schemes for stochastic volatility models. *Quantitative Finance*, 6(6):513–536, 2006.

[31] P. E. Kloeden, E. Platen, P. E. Kloeden, and E. Platen. *Stochastic Differential Equations*. Springer, 1992.

[32] S. G. Kou. A jump-diffusion model for option pricing. *Management science*, 48(8): 1086–1101, 2002.

[33] D. B. Madan and E. Seneta. The variance gamma (VG) model for share market returns. *Journal of business*, pages 511–524, 1990.

[34] R. C. Merton. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, pages 141–183, 1973.

[35] R. C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1-2):125–144, 1976.

[36] N. Metropolis and S. Ulam. The monte carlo method. *Journal of the American Statistical Association*, 44(247):335–341, 1949.

[37] M. Rubinstein. Implied binomial trees. *The Journal of Finance*, 49(3):771–818, 1994.

[38] L. O. Scott. Option pricing when the variance changes randomly: Theory, estimation, and an application. *Journal of Financial and Quantitative Analysis*, 22(4):419–438, 1988.

[39] P. Tankov. *Financial modelling with jump processes*. CRC press, 2003.

[40] Y. Zhu and J. E. Zhang. Variance term structure and VIX futures pricing. *International Journal of Theoretical and Applied Finance*, 10(01):111–127, 2007.

[41] G. Zumbach. Time reversal invariance in finance. *Quantitative Finance*, 9(5):505–515, 2009.

[42] G. Zumbach. Volatility conditional on price trends. *Quantitative Finance*, 10(4): 431–442, 2010.

A | Appendix A

In the case of a stochastic process $(Y_t)\, t \geq 0$ with trajectories exhibiting bounded variation and a sufficiently smooth function f, the integral of $Zt = f(Y_t)$ with respect to dY_t can be defined using the Riemann[1] integral:

$$\int_0^t Z_s dY_s = \int_0^t f(Y_s)\, dY_s \tag{A.1}$$

However, if the trajectories of $(Y_t)_{t\geq 0}$ lack bounded variation and the function f lacks sufficient regularity, the Riemann integral may not be well-defined. This is particularly true for Brownian motion, where the integral

$$\int f(W_s)\, dW_s \tag{A.2}$$

cannot generally be defined as a Riemann integral. To overcome this limitation, Itô's stochastic integral is introduced, represented as $\int_0^t H_s dW_s$ for processes $(H_t)\, t \geq 0$ that meet a "suitable measurability assumption" and satisfy $P\left(\int 0^t H_s^2 ds < +\infty\right) = 1$. Under these conditions, the following equation can be expressed:

$$Y_t = Y_0 + \int_0^t g(Y_s, s)\, ds + \int_0^t \sigma(Y_s, s)\, dW_s \tag{A.3}$$

Here, the first integral is interpreted as a classic Riemann integral (under suitable regularity assumptions on g and σ), while the second is an Itô integral. This expression can also be represented as a Stochastic Differential Equation (SDE):

$$\begin{cases} dY_t = g(Y_t, t)\, dt + \sigma(Y_t, t)\, dW_t \\ Y_0 = y \end{cases} \tag{A.4}$$

[1]The definite integral according to Riemann is a mathematical operator that associates real functions of real variable with the area subtended by the graph on an interval of your choice, under appropriate assumptions.

It is important to note that both formulations are equivalent; the latter is simply an alternative representation of the former.

A.1. Itô formula

An essential result in the field of stochastic analysis is the well-known Itô Lemma. In what follows, we will denote by $f(x,t)$ in $C^{2,1}(\mathbb{R},\mathbb{R}^+)$ a function that is continuously differentiable twice in x and once in t. Similarly for $f(x_1,\ldots,x_n,t)$ in $C^{2,1}(\mathbb{R}^n,\mathbb{R}^+)$.

Itô's Lemma

Let $f(x,t) \in C^{2,1}(\mathbb{R},\mathbb{R}^+)$be a given function. If the process $(X_t)_{t\geq 0}$ satisfies the following equation:

$$dX_t = \mu(t)dt + \sigma(t)dW_t \tag{A.5}$$

then the process $(Z_t)_{t\geq 0}$, defined by $Z_t := f(X_t,t)$, satisfies

$$\begin{aligned} dZ_t &= \frac{\partial f}{\partial t}(X_t,t)\,dt + \frac{\partial f}{\partial x}(X_t,t)\,dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(X_t,t)\,\sigma^2(t)dt \\ &= \left(\frac{\partial f}{\partial t} + \mu\frac{\partial f}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2}\right)(X_t,t)\,dt + \sigma(t)\frac{\partial f}{\partial x}(X_t,t)\,dW_t \end{aligned} \tag{A.6}$$

This formula will often be applied in different contexts in this Thesis.

A.2. Itô formula: B&S application

Let us consider the Black & Scholes model, recalling Eq. (1.2) in which the model was defined as

$$\begin{cases} dS_t = rS_t dt + \sigma S_t dW_t \\ S_0 \end{cases}$$

We apply Itô's formula to obtain Eq. (A.7) as follows

$$S_t = S_0 e^{\left(r-\frac{1}{2}\sigma^2\right)t+\sigma W_t} \tag{A.7}$$

where r is the interest rate, σ the variance of the underlying, and W_t the Brownian motion (Wiener process). We choose $f(x) = \ln x$, consequently we deduce that $f'(x) = \frac{1}{x}$ and

$f'(x) = -\frac{1}{x^2}$. Itô's formula implies that:

$$\begin{aligned} df(S_t) = d(\ln S_t) &= f'(S_t)\, dS_t + \frac{1}{2} f''(S_t) [\sigma S_t]^2 dt \\ &= \frac{1}{S_t} dS_t - \frac{1}{2} \frac{1}{(S_t)^2} [\sigma S_t]^2 dt \\ &= \frac{1}{S_t} [r S_t dt + \sigma S_t dW_t] - \frac{1}{2} \sigma^2 dt \\ &= \left(r - \frac{1}{2} \sigma^2 \right) dt + \sigma dW_t. \end{aligned} \tag{A.8}$$

It follows that, considering the initial condition S_0, we can solve the SDE as

$$\ln(S_t) = \ln(S_0) + \left(r - \frac{1}{2} \sigma^2 \right) t + \sigma W_t \tag{A.9}$$

hence

$$S_t = S_0 e^{\left(r - \frac{1}{2}\sigma^2\right) t + \sigma W_t} \tag{A.10}$$

By defining the log-price ($\ln(S_t)$), as X_t, we can define its following SDE:

$$\begin{cases} dX_t = \left(r - \frac{1}{2} \sigma^2 \right) dt + \sigma dW_t \\ X_0 = \ln(S_0) \end{cases} \tag{A.11}$$

hence

$$X_t = X_0 + \left(r - \frac{1}{2} \sigma^2 \right) t + \sigma W_t \tag{A.12}$$

List of Figures

List of Tables