Implementing and Calibrating the Regime-Switching Black-Scholes Model

Student name: O.H. Valmas Supervisor name: A. Wade

Submitted as part of the degree of MSc in Scientific Computing and Data Analysis to the Board of Examiners in the Department of Computer Sciences, Durham University

***Abstract* —** The shortcomings of the Black-Scholes model in the context of pricing European options are well understood. The assumption that the volatility of the underlying is constant means that simple Black-Scholes models cannot accurately price options in practice. This paper uses the regime-switching Black-Scholes approach to model European call options using both analytic solutions and Monte Carlo methods in order to replicate the variable volatility present in the market. We demonstrate that the model can produce the characteristic volatility smiles and volatility term structures with arbitrary parameters. The price predicted by the model is calibrated to the market price of SPY call options using least squares minimisation. The optimised parameters are given and the calibrated prices and implied volatility surface is plotted. We also investigate the cause of some of the errors for options with certain expiries and moneyness. Finally, we evaluate the work, suggest some improvements and highlight further research directions in the literature and propose some of our own.

***Keywords* —** regime-switching, Black-Scholes, option pricing, calibration, stochastic volatility, continuous-time Markov chains, volatility surface, least squares minimisation

Contents

I Introduction **2**

II Related Work **3**
- A Black-Scholes Pricing of Financial Derivatives . . . 3
- B Shortcomings of the Black-Scholes Model . . . 3
- C The Volatility Surface . . . 3
- D Stochastic Volatility . . . 4
- E Continuous-time Markov Chains . . . 5
- F Regime-Switching Models . . . 7

III Solution **10**
- A Simulating Stochastic Differential Equations . . . 10
- B Implementing Monte-Carlo Pricing . . . 10
- C Obtaining Market Data . . . 12
- D Determining the Risk-free Rate . . . 12
- E Calibration of the Model . . . 12
- F Calculating the Implied Volatility . . . 13

IV Results and Evaluation **14**
- A Verification of the Continuous-Time Markov Chain . . . 14
- B Comparing Analytical and Monte-Carlo Prices . . . 15
- C Volatility Smile and Term Structure . . . 16
- D Price Calibration . . . 16

V Conclusion **19**

I INTRODUCTION

It is well known that the volatility of risky assets, such as common stock, changes over time [1]. Empirical studies have established various stylised features of the volatility [2]. These features include volatility clustering, leverage effects, increased volatility following non-trading periods, higher volatility after the releases of important information such as earnings reports, and co-movements in volatility across distinct assets. Therefore, various models of changing volatility have been developed to capture these phenomena. Three classes of models that have received considerable industry attention are ARCH type models, stochastic volatility models, and regime-switching models.

It is widely accept that the classical Black–Scholes model with constant volatility does not fully reflect the stochastic nature of financial markets [3]. Consequently, there is a need for more realistic models that better reflect random market movements. One such formulation is a model with regime-switching, where the key parameters of an asset depend on a market regime that switches among a finite set of states. Since being introduced by Hamilton [4], growing empirical evidence has suggested that the distributions of asset returns are, in some cases, better described by a regime-switching process [5, 6]. From an economic perspective, regime-switching behavior captures the ever-changing preferences and beliefs of market participants surrounding asset prices as the state of the underlying financial market changes. It therefore makes intuitive sense to implement into a model.

This paper investigates the efficient implementation and calibration of a regime-switching Black-Scholes model. We first discuss the implementation of a continuous-time Markov chain. We verify our implementation by evaluating the accuracy of the emergent stationary distribution and providing some sample outputs. We then incorporate the Markov chain into the Black-Scholes model to simulate the price of an asset during regime-switches. We derive and implement the analytical solutions from the literature to calculate the price of call options under a two-state regime-switching model. Monte Carlo methods are then employed to numerically calculate the same option price and verify the analytical implementations. These implementations allow us to explore the properties of the regime-switching model and determine whether it generates the implied volatility smiles and term structures found in real financial markets. Finally, we carry out the calibration of the regime-switching model to real option prices and demonstrate the reconstruction of the implied volatility surface found in SPY options. Throughout the work, SPY refers to an ETF (exchange-traded fund) that tracks the S&P 500 index, a stock market index measuring the performance of 500 large market capitalisation publicly traded companies in the United States.

II RELATED WORK

A Black-Scholes Pricing of Financial Derivatives

The Black-Scholes model [7], introduced 50 years ago, revolutionized the field of quantitative finance and continues to have a profound impact on the field. This groundbreaking model transformed pricing options, instruments which were previously considered challenging to value. By more accurately pricing financial products and offering theoretical solutions under its assumptions, the Black-Scholes model has since played a crucial role in pricing derivatives. Building on the work of Black and Scholes, Merton further enhanced the usability of the Black-Scholes Equation by presenting an alternative derivation of the Black-Scholes formula under weaker assumptions such as dynamic interest rates[8]. In the Black-Scholes model, the price of a risky asset S_t is described by the stochastic differential equation:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \qquad \mu, \sigma \in \mathbb{R} \tag{1}$$

where μ is the annualised rate of return of S, σ is the standard deviation of the asset's returns, and W_t is a Wiener process. The price of a risk-free asset B_t, commonly referred to as cash or bonds, is described by the ordinary differential equation:

$$dB_t = rB_t dt \qquad r \in \mathbb{R} \tag{2}$$

where r is the annualised risk-free interest rate. Under the Black-Scholes model, several assumptions about the assets are made. It is assumed that the the risk-free interest rate r, drift μ and volatility σ are all constant. Additionally, the use of geometric Brownian motion to model the price process of the risky asset assumes its returns are log-normally distributed. The model also assumes that the risky asset does not pay a dividend. Furthermore, the model makes several assumptions about the market itself. It assumes that there exists no arbitrage in the market, and that a market participant can freely buy and sell any amount of stock or cash, including fractional amounts. Finally, it assumes that there are no transaction costs, slippage, taxes or restrictions on short selling.

B Shortcomings of the Black-Scholes Model

While the Black-Scholes model has been instrumental in option pricing, it does have several shortcomings that must be acknowledged and that motivate this work. First of all, the assumptions discussed above often do not hold true in real-world financial markets. These assumptions oversimplify the complexity of the markets, resulting in potential inaccuracies during pricing. The model further assumes that the returns on the underlying asset follow a log-normal distribution. Empirical evidence suggests however that asset returns often exhibit fat tails and experience more extreme movements than what a log-normal distribution predicts [9]. This discrepancy can result in underestimating the likelihood of positive extreme events, such as company acquisition, and of negative extreme events such as poor earnings reports and market crashes.

The Black-Scholes model assumes the existence of a risk-free asset and the ability to continuously trade the underlying asset. However, in reality, markets can be incomplete, with limited availability of, or trading restrictions on, certain assets. The model does not account for various market frictions, such as liquidity constraints for less popular products, and market impact. In reality, these frictions can noticeably affect option prices. Finally, the model assumes constant volatility of the risky assert, which motivates using variable volatility models to augment the vanilla Black-Scholes in order to capture the volatility dynamics observed in markets.

C The Volatility Surface

The Black-Scholes, while an elegant model, does not perform very well in practice; it is well known that stock prices occasionally jump and do not always move in the smooth manner predicted by geometric Brownian motion [10, 11]. Stock prices tend also to have fatter tails than those predicted by geometric Brownian motion. Most importantly for this paper, if the Black-Scholes model were correct then we should have a flat implied volatility surface. The volatility function $\sigma(K, T)$ is defined as

$$C(K, T) := BS(T, r, K, \sigma(K, T)) \tag{3}$$

where $C(K,T)$ denotes the current market price of a call option with time-to-maturity T and strike K, and $BS(T,r,K,\sigma)$ is the Black-Scholes formula for pricing a European call option under volatility σ. It is noted that all formulae omit the current price of the underlying, S, as this is kept constant. We recognise here that put-call parity states that the price of a put and a call with the same strike and expiration must have the same implied volatility [12]. This means that either the price of a put option or a call option can be used to calculate implied volatility and will lead to the same result. Put another way, $\sigma(K,T)$ is the volatility that, when substituted back into the Black-Scholes formula, gives the market price, $C(K,T)$. As the Black-Scholes formula is continuous and increasing in σ, there will always be a unique solution, $\sigma(K,T)$ [13]. If the Black-Scholes model were correct then the volatility surface would be flat with $\sigma(K,T) = \sigma$ for all K and T. In practice, however, not only is the volatility surface not flat but it actually varies, often significantly, with time.

There are several plots used throughout the literature to characterise the non-constant volatility present in real markets. The volatility smile is a plot of the implied volatility of an option with a certain expiration as a function of its strike price. The volatility term structure is a plot of the implied volatility of an option with a given strike as a function of its expiry. A three dimensional plot of the implied volatility as a function of both the strike price and the time to maturity is known as a volatility surface.

Depending on the shape of the volatility smile, two variants can be distinguished: the smile and the smirk [14]. The volatility smile shape can be frequently observed in near-term equity options and options in the foreign exchange market. Volatility smile patterns indicate that demand is larger for options that are deeply in-the-money or out-of-the-money. The volatility smirk, in contrast, has two flavours: the forward skew and the reverse skew. While the forward skew shape typically appears for options in the commodities market, the reverse skew shape usually occurs in longer-term equity options and index options. The implied volatility for options in the reverse skew shape increases with lower strike prices and decreases with higher strikes prices. This, in turn, suggests that out-of-the money (OTM) calls and in-the-money (ITM) puts are discounted relative to ITM calls and OTM puts, indicating a lower demand in the market. On the other hand, the implied volatility for options in the forward skew shape decreases with lower strikes and increases with higher strikes. This suggests that ITM calls and OTM puts have a lower demand than OTM calls and ITM puts.

Clearly in order to generate implied volatility smiles and surfaces, we must explore how to calculate the implied volatility itself. Ideally, calculating the implied volatility would be as simple as finding some function $BS^{-1}(C(S,K,T),\cdot)$ that is the inverse of the Black-Scholes formula. However, a closed-form solution for determining the implied volatility does not exist for the vanilla Black-Scholes formula [15]. Instead, root-finding algorithm is commonly used to find a solution σ_i to the following equation

$$BS(T,r,K,\sigma_i) - C(K,T) = 0 \tag{4}$$

where $C(K,T)$ is the current market price of the European call option. A popular choice for finding this root is the Newton-Raphson method where σ_i is calculated iteratively. In the case of the Black-Scholes formula, we use the partial derivative with respect to σ. This derivative is known as Vega (ν) and is one of the option *Greeks*, a collection of financial measures of an option's sensitivity to its underlying parameters. σ_i is calculated iteratively using the following equations:

$$\sigma_{i+1} = \sigma_i - \frac{BS(T,r,K,\sigma_i) - C(K,T)}{\nu(\sigma_i)}, \qquad \nu = \frac{\partial BS}{\partial \sigma} = S\sqrt{T}\phi(d_1), \qquad d_1 = \frac{\ln(\frac{S}{K}) + (r + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}. \tag{5}$$

The iterative process continues until we reach the desired degree of accuracy and the error is within some tolerance ϵ:

$$|BS(\cdot) - C(K,T)| \leq \epsilon, \epsilon \in \mathbb{R}. \tag{6}$$

D Stochastic Volatility

Incorporating stochastic volatility into the Black-Scholes model recognises the fact that volatility is not a constant parameter over time as assumed. This allows us to capture the time-varying nature of market volatility more effectively. The Black-Scholes model also assumes constant volatility across all strike prices at any given time, disregarding the observed volatility smile observed in real option markets.

Several approaches have been proposed to tackle stochastic volatility which involve replacing the constant σ with a stochastic process $\nu(t)$. Heston [16] proposed a model where the underlying asset price follows geometric Brownian motion, as in Black-Scholes. However, the volatility is assumed to follow a Cox–Ingersoll–Ross process, capturing the mean-reverting nature of volatility observed in financial markets. As such, the risky asset is defined by the system of stochastic differential equations

$$dS_t = \mu S_t dt + \sqrt{\nu(t)} S_t \, dW_t^S \tag{7}$$

$$d\nu(t) = \kappa[\theta - \nu(t)]dt + \sigma\sqrt{\nu(t)} \, dW_t^\nu \tag{8}$$

$$dW_t^S dW_t^\nu = \rho \, dt \tag{9}$$

where W_t^S and W_t^ν are both standard Wiener processes. Heston's model has five parameters: the initial variance ν_0, the long run variance θ, the correlation coefficient of the two Wiener processes ρ, the rate of mean reversion κ and the volatility of the volatility σ. The ρ parameter allows the model to capture the relationship between changes in the underlying asset returns and changes in volatility. Fouque et al. [5, 6] propose a similar approach to handling stochastic volatility but instead propose the use of an Ornstein-Uhlenbeck process to model the instantaneous volatility:

$$dS_t = \mu S_t dt + f(\nu(t)) S_t \, dW_t \tag{10}$$

$$d\nu(t) = \alpha[m - \nu(t)] \, dt + \beta \, d\hat{Z}_t \tag{11}$$

$$\hat{Z}_t = \rho W_t + \sqrt{1 - \rho^2} Z_t \tag{12}$$

where p is the correlation between W_t and $\hat{Z}_t$.

E Continuous-time Markov Chains

A continuous-time Markov chain is a continuous stochastic process where, for each state, the process will change state according to an exponential random variable and then move to a different state as specified by the probabilities of a stochastic matrix. An equivalent formulation describes the process as changing state according to the least value of a set of exponential random variables, one for each possible state to which it can move, with the parameters for these random variables determined by the current state. A continuous-time Markov chain $X(t)$ is defined by two components: a jump chain and a set of holding time parameters λ_i. The jump chain consists of a countable set of states $S \subset \{1, 2, ...\}$ along with transition probabilities $p_{i,j}$. It is first assumed that $p_{i,i} = 0$ for all non-absorbing states $i \in S$. We also assume that

1. if $X(t) = i$, the time until the state changes is determined by an Exponential distribution Exp(λ_i), and
2. if $X(t) = i$, the next state will be state j with probability p_{ij}.

In the interest of completeness and because it is used later, we define here the exponential distribution. The probability density function of an exponential distribution, where $\lambda > 0$ is the rate parameter of the distribution, is

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0. \end{cases} \tag{13}$$

The continuous-time Markov chain $X(t)$ satisfies the Markov property, the assertion that the future evolution of the chain in the interval between t_n and t_{n+1}, given the present, is independent of the past. That is, for all $t_0 < t_1 < ... < t_n < t_{n+1}$, we have

$$P(X(t_{n+1}) = j | X(t_n) = i, X(t_{n-1}) = i_{n-1}, ..., X(t_1) = i_1) = P(X(t_{n+1}) = j | X(t_n) = i). \tag{14}$$

Thus, the transition probability $p_{ij}(t)$ is defined as

$$\begin{aligned} p_{i,j}(t) &= p(X(t+s) = j | X(s) = i) \\ &= p(X(t) = j | X(0) = i) \end{aligned} \tag{15}$$

for all $s, t \in [0, \infty)$. This allows us to construct the transition matrix $P(t)$ which, for r states $1, 2, .., r$, is defined as

$$P(t) = \begin{bmatrix} p_{11}(t) & p_{12}(t) & \cdots & p_{1r}(t) \\ p_{21}(t) & p_{22}(t) & \cdots & p_{2r}(t) \\ \vdots & \vdots & \ddots & \vdots \\ p_{r1}(t) & p_{r2}(t) & \cdots & p_{rr}(t) \end{bmatrix} \tag{16}$$

There are various approaches for calculating the transition matrix from a given generator matrix Q. It is known that $P(t)$ defined above satisfies the Komologrov forward equation:

$$P'(t) = P(t)Q. \tag{17}$$

The solution to this forward equation can be recognised as $P(t) = e^{tQ}$ [17]. The matrix exponential can be solved elegantly by diagonalising the generator matrix Q. Consider, without loss of generality, a 2-state continuous time Markov chain with holding time parameters $\lambda_1 = \alpha$ and $\lambda_2 = \beta$. The Q matrix is therefore defined as

$$Q = \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix}. \tag{18}$$

Suppose that Q has two distinct eigenvectors. Then, letting D be a diagonal matrix consisting of the eigenvalues of Q, we can decompose Q into

$$Q = XDX^{-1} \tag{19}$$

where X consists of the eigenvectors of Q (ordered similarly to the order of the eigenvalues in D). We can diagonalise Q as follows:

$$X = \begin{bmatrix} 1 & \frac{-\alpha}{\beta} \\ 1 & 1 \end{bmatrix}, D = \begin{bmatrix} 0 & 0 \\ 0 & -(\alpha+\beta) \end{bmatrix}, X^{-1} = \begin{bmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ -\frac{\beta}{\alpha+\beta} & \frac{\beta}{\alpha+\beta} \end{bmatrix} \tag{20}$$

Using the matrix exponential, we get that

$$e^{Qt} = \sum_{n=0}^{\infty} \frac{t^n (XDX^{-1})^n}{n!} = P(\sum_{n=0}^{\infty} \frac{t^n D^n}{n!}) = Xe^{Dt}X^{-1}, \tag{21}$$

where, because D is diagonal, e^{Dt} is also diagonal with diagonal elements $e^{\eta_i t}$ where η_i is the i-th eigenvalue. Hence, $P(t) = e^{tQ}$ can be calculated using the diagonalisation above as

$$P(t) = e^{tQ} = \begin{bmatrix} 1 & \frac{-\alpha}{\beta} \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & e^{-(\alpha+\beta)t} \end{bmatrix} \cdot \begin{bmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ -\frac{\beta}{\alpha+\beta} & \frac{\beta}{\alpha+\beta} \end{bmatrix} = \begin{bmatrix} \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}e^{-(\alpha+\beta)t} & \frac{\alpha}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta}e^{-(\alpha+\beta)t} \\ \frac{\beta}{\alpha+\beta} - \frac{\beta}{\alpha+\beta}e^{-(\alpha+\beta)t} & \frac{\alpha}{\alpha+\beta} + \frac{\beta}{\alpha+\beta}e^{-(\alpha+\beta)t} \end{bmatrix}. \tag{22}$$

If we consider the limit of $P(t)$ as $t \to \infty$, we arrive at a key concept in Markov chains: the stationary distribution. The stationary distribution for an irreducible, recurrent continuous-time Markov chain is the probability distribution to which the process converges for large values of t. Hence, for a the continuous-time Markov chain discussed above, the stationary distribution is defined as

$$\lim_{t\to\infty} P(t) = \lim_{t\to\infty} \begin{bmatrix} \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}e^{-(\alpha+\beta)t} & \frac{\alpha}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta}e^{-(\alpha+\beta)t} \\ \frac{\beta}{\alpha+\beta} - \frac{\beta}{\alpha+\beta}e^{-(\alpha+\beta)t} & \frac{\alpha}{\alpha+\beta} + \frac{\beta}{\alpha+\beta}e^{-(\alpha+\beta)t} \end{bmatrix} = \begin{bmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{bmatrix}. \tag{23}$$

We observe that each row is identical as $t \to \infty$ as the distribution does not depend on the starting state. The stationary distribution is commonly denoted as the row vector $\pi = (\pi_1, \pi_2)$. This row vector π can be calculated simply by solving the below equation with accompanying constraint for any square generator matrix Q:

$$\pi Q = 0, \quad \sum_{i \in S} \pi_i = 1. \tag{24}$$

The stationary distribution of a continuous-time Markov chain will prove useful later on as it will allow us to calculate the price of options as a weighted average according to this distribution, as proposed by Fuh et al. [18].

F Regime-Switching Models

Another approach to stochastic volatility is regime-switching models, such as Markovian regime-switching, dating back at least to Hamilton [4, 19]. Markov switching models have been successfully implemented in [20] where a hidden Markov process is used to represent the state of information in the investors' community. Fuh et al. [18] consider a model of time-varying volatility to generalise Black-Scholes to incorporate regime-switching properties. Similarly to the above formulation, they use the state of this hidden Markov process to determine instantaneous drift and volatilities during simulation. Naik [21] provides a closed-form formula for the arbitrage-free price of European call options, provided the Markov process has a finite number of states. Sohbrai [22] implements this closed-form formula to price European call options in a regime-switching environment exactly. The author also compares the exact prices to prices determined both by Monte-Carlo simulations and a novel, faster method involving the inverse Fourier transform.

Regime-switching Black-Scholes models integrate variable volatility by introducing a continuous-time Markov chain η_t which keeps track of the state of the market at time t. Then, the original Black-Scholes SDE is reformulated accordingly, although the risk-free asset is unchanged:

$$dS_t = \mu_{\eta_t} S_t dt + \sigma_{\eta_t} S_t dW_t \qquad \mu_{\eta_t}, \sigma_{\eta_t} \in \mathbb{R} \tag{25}$$

where $(\eta_t, t \geq 0)$ and $(W_t, t \geq 0)$ are independent and $(W_t, t \geq 0)$ is standard Wiener process. $(\eta_t, t \geq 0)$ is a continuous-time Markov chain on state space $S = \{1, 2, ..., M\}$ where M is the number of states. The Markov chain also has an associated $M \times M$ generator matrix Q which determines the dynamics of the chain. For a continuous-time Markov chain with M states, Q can be written generally as

$$Q = \begin{bmatrix} \lambda_{1i} & \lambda_{12} & \cdots & \lambda_{1M} \\ \lambda_{21} & \lambda_{2i} & \cdots & \lambda_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{M1} & \lambda_{M2} & \cdots & \lambda_{Mi} \end{bmatrix} \tag{26}$$

where $\lambda_{kl} \geq 0$ for $k \neq l$ and $\sum_{l=1}^{M} \lambda_{kl} = 0$ for $k = 1, \ldots, M$. In the simplest case, there exists only two states, $\eta_t \in \{1, 2\}$, which could be thought of representing a bull market, a period in which prices of assets are expected to rise, and a bear market, a period when prices are expected to fall. Clearly, if $\mu_1 = \mu_2$ and $\sigma_1 = \sigma_2$, the regime-switching model will replicate the regular Black-Scholes model exactly.

Although the pricing formula of European put and call options using the vanilla Black-Scholes equation is ubiquitous, a closed-form solution for the price of an option under a regime-switching model is not. Naik [21] derives option pricing equations under the assumption that, at the instant the volatility of the underlying process switches, there is no change in the price of the asset. In other words, the price process of the risk asset has continuous sample paths. These equations will form the basis of the analysis of the analytic solution to pricing European options throughout this work.

Let $C_k(S, T)$ be the value of a European option with regime k, strike price K and expiry date T. The call option price $C_k(S, T)$ in each regime k satisfies the following equations

$$\begin{gathered} \frac{\partial C_k}{\partial t} + \frac{1}{2}\sigma_k^2 S^2 \frac{\partial^2 C_k}{\partial S^2} + r_k S \frac{\partial C_k}{\partial S} - r_k C_k + \sum_{\substack{k,l=1 \\ k \neq l}}^{M} \lambda_{kl}(C_l - C_k) = 0 \\ C_k(S, T) = \max(S - K, 0), k = 1, \ldots, M, \end{gathered} \tag{27}$$

where σ_k and r_k are the volatility and risk-free interest rate for regime k, λ_{kl} are the transition rates from regime k to regime l from the generator matrix Q defined above. As above, consider the simplified case where the continuous-time Markov chain has $M = 2$ states and $\sigma_1 > \sigma_2$. The price of this European call option when $\eta_t = 1$ is

$$\begin{gathered} C_1 := C(S, \sigma_1, t) = \int_0^{T-t} C^{BS}\left(S, K, r, T - t, \sqrt{\frac{s(x)}{T-t}}\right) f(x|\sigma_1)\, dx, \\ s(x) = (\sigma_1)^2 x + (\sigma_2)^2 (T - t - x) \quad 0 \leq x \leq T - t, \end{gathered} \tag{28}$$

where $C_{BS}(\cdot)$ is the vanilla Black-Scholes formula for a European call option, and $f(x|\sigma_1)$ denotes the conditional density for the occupation time of the volatility process in state σ_1. The density is given by

$$f(x|\sigma_1) = e^{-\lambda_{12}x-\lambda_{21}(T-t-x)}\Big(\delta_0(T-t-x) + g_1(x)I_1(2(h(x)) + \lambda_{12}I_0(2(h(x))\Big), \tag{29}$$

where $\delta_0(x)$ is the Dirac delta function, $I_p(x)$ is the modified Bessel function of order p,

$$I_p(x) = \sum_{k=0}^{\infty} \frac{1}{k!\Gamma(k+p+1)}(\frac{x}{2})^{2k+p} = (\frac{x}{2})^p \sum_{k=0}^{\infty} \frac{1}{k!(k+p)!}(\frac{x}{2})^{2k}, \tag{30}$$

and

$$h(x) = \sqrt{\lambda_{12}\lambda_{21}x(T-t-x)}, \quad g_1(x) = \sqrt{\frac{\lambda_{12}\lambda_{21}x}{(T-t-x)}}. \tag{31}$$

The price C_2 is the same as C_1 except that the conditional density $f(x|\sigma_2)$ is defined as

$$f(x|\sigma_2) = e^{-\lambda_{12}x-\lambda_{21}(T-t-x)}\Big(\delta_0(x) + g_2(x)I_1(2(h(x)) + \lambda_{21}I_0(2(h(x))\Big), \tag{32}$$

with

$$g_2(x) = \sqrt{\frac{\lambda_{12}\lambda_{21}(T-t-x)}{x}}. \tag{33}$$

Fuh et al. [18] provide a rigorous proof of the above in their appendix. Evaluating the price of European call options using these equations is non-trivial due to the presence of a double integral, as the vanilla Black-Scholes itself uses the normal distribution cumulative density function. This motivates some transformations:

$$\begin{aligned}
C_1 &= \int_0^{T-t} C_{BS}\left(\cdot, \sqrt{\frac{s(x)}{T-t}}\right) e^{-\lambda_{12}x-\lambda_{21}(T-t-x)}\Big(\delta_0(T-t-x) + g_1(x)I_1(2(h(x)) + \lambda_{12}I_0(2(h(x))\Big)\, dx \\
&= \underbrace{\int_0^{T-t} C_{BS}\left(\cdot, \sqrt{\frac{s(x)}{T-t}}\right) e^{-\lambda_{12}x-\lambda_{21}(T-t-x)}\delta_0(T-t-x)\, dx}_{\text{A}} \\
&+ \underbrace{\int_0^{T-t} C_{BS}\left(\cdot, \sqrt{\frac{s(x)}{T-t}}\right) e^{-\lambda_{12}x-\lambda_{21}(T-t-x)}\Big(g_1(x)I_1(2(h(x)) + \lambda_{12}I_0(2(h(x))\Big)\, dx}_{\text{B}}.
\end{aligned}$$

We now introduce two useful properties of the Dirac delta function:

$$\int_{-\infty}^{+\infty} f(t)\delta(t-T)\, dt = f(T) \quad \text{and} \quad \delta(-x) = \delta(x). \tag{34}$$

The first of these properties is the integral of the time-delayed Dirac delta, and is commonly called the *sifting property*. Using these two properties, we can express A in terms of a Black-Scholes call option price:

$$\begin{aligned}
A &= \int_0^{T-t} C_{BS}\left(S, K, r, T-t, \sqrt{\frac{s(x)}{T-t}}\right) e^{-\lambda_{12}x-\lambda_{21}(T-t-x)}\delta_0(x-(T-t))\, dx \\
&= C_{BS}\left(S, K, r, T-t, \sqrt{\frac{s(T-t)}{T-t}}\right) e^{-\lambda_{12}x-\lambda_{21}(T-t-(T-t))} \\
&= C_{BS}\Big(S, K, r, T-t, \sigma_1\Big) e^{-\lambda_{12}(T-t)-\lambda_{21}(T-t-(T-t))} \\
&= C_{BS}\Big(S, K, r, T-t, \sigma_1\Big) e^{-\lambda_{12}(T-t)}.
\end{aligned}$$

This leaves us with only B to evaluate numerically. Consider the behaviour of $g_1(x)I_1(2(h(x))+\lambda_{12}I_0(2(h(x))$ as $x \to T-t$. As $x \to T-t$, $g_1(x) \to +\infty$. However, $h(T-t) = 0 \implies I_1(0) = 0$. This makes the integrand

of B indeterminate at the upper bound on integration i.e. when $x = T - t$. This is solved by evaluating the following limit using the definition of the modified Bessel function:

$$\lim_{x \to T-t} g_1(x) I_1(2h(x)) = \lim_{x \to T-t} \sqrt{\lambda_{12}\lambda_{21}} \frac{\sqrt{x}}{\sqrt{T-t-x}} \left(\sum_{m=0}^{\infty} \frac{1}{m!(m+1)!} \sqrt{\lambda_{12}\lambda_{21} x (T-t-x)}^{2m+1} \right) \tag{35}$$

$$= \lambda_{12}\lambda_{21}(T-t). \tag{36}$$

These transformations are applied similarly to $f(x|\sigma_2)$, enabling the numerical calculation of both C_1 and C_2. Now that we have closed-form solutions for C_1 and C_2, they are combined using the weighted sum proposed by Fuh et al. [18]

$$C = \pi_1 C_1 + \pi_2 C_2 \tag{37}$$

where $\pi = (\pi_1, \pi_2)$ is the stationary distribution of the continuous-time Markov chain.

III SOLUTION

A Simulating Stochastic Differential Equations

As established previously, the price of a risky asset S_t is described by geometric Brownian motion and can be solved analytically using Itô's Lemma [23]:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \qquad S_t = S_0 \exp\left((\mu - \frac{\sigma^2}{2})t + \sigma W_t\right). \tag{38}$$

A solution to a stochastic differential equation is itself a stochastic function, which means that its value S_t at any given time t is a random variable. We first simulate four sample paths for S_t to illustrate this point and plot them in Figure 1. Sample paths differ because of different realizations of this Brownian motion term and from different values of μ and σ. The time interval $[0, T]$ is first divided into N distinct chunks $0 = t_0 < t_1 < ... < t_N = T$ and S_t is computed at each point t_i. To compute the value of W_t at those points, we construct a sample path of increments $\Delta W_{t_n} = W_t - W_{t_{n-1}}$. We know that each increment is distributed according to $\sqrt{\Delta t} \cdot N(0, 1)$, where $\Delta t = \frac{T}{N}$. We then create the sample path of Brownian motion B by summing the increments, and compute the exact solution S_t using the analytical solution above. The above

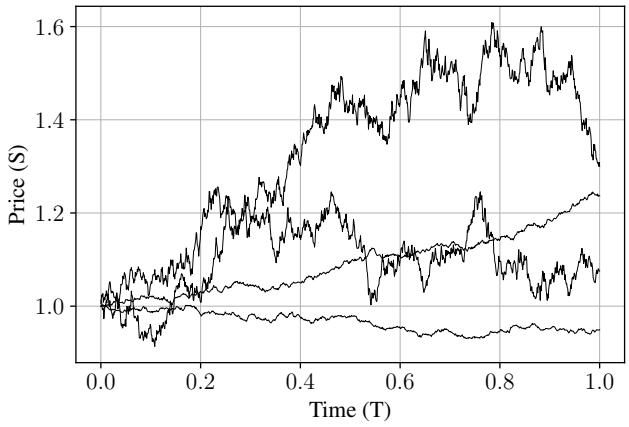


Figure 1: Sample paths of a geometric Brownian motion process

shows that the value S_t for a given t differs across sample paths (because it is a random variable). In order to demonstrate how the sample paths vary over time under a regime-switching process, we first must simulate the underlying continuous-time Markov chain. This is achieved using *Gillespie's algorithm* and allows us to illustrate the evolution of the continuous-time Markov chain graphically. Now that we have a sample trajectory of the Markov chain, we can use it to generate paths under the regime-switching process. Figure 2 is an example of one of these paths under a 3-state continuous-time Markov chain, where the three regimes have annualised rates of return $\mu_1 = 0.3, \mu_2 = -0.3, \mu_3 = 0.1$ and volatilities $\sigma_1 = 0.1, \sigma_2 = 0.1, \sigma_2 = 0.2$.

B Implementing Monte-Carlo Pricing

Now that we can simulate the price our risky asset S_t under the regime-switching process, we can use the Monte-Carlo technique to price European options using our model. We again assume a 2-state model. In order to price non-path dependent options such as European call options, we require only the final stock price at maturity S_T, rather than the stock price trajectory $\{S_t\}_{t\in[0,T]}$. This allows us to optimise the Monte Carlo scheme significantly. We can therefore implement the following algorithm for obtaining pseudo-random samples from the terminal stock price distribution under the risk-neutral measure.

Consider the trajectory of the Markov chain $\{\eta_t\}_{t\in[0,T]}$, conditional on $\eta_0 = 1$. We know that the amount of time (beginning at $t = 0$) until the Markov chain leaves state 1 has an exponential distribution with rate λ_1. Let $t_1, t_2, t_3 \in [0, T]$ denote successive jump times between states. Hence, $t_1 \sim \text{Exp}(\lambda_1)$ and $P(t = t_1) = \lambda_1 e^{-\lambda_1 t_1}$. We draw a uniform random variable u on the interval $[0, 1]$ to simulate the probability $P(t = t_1)$ as

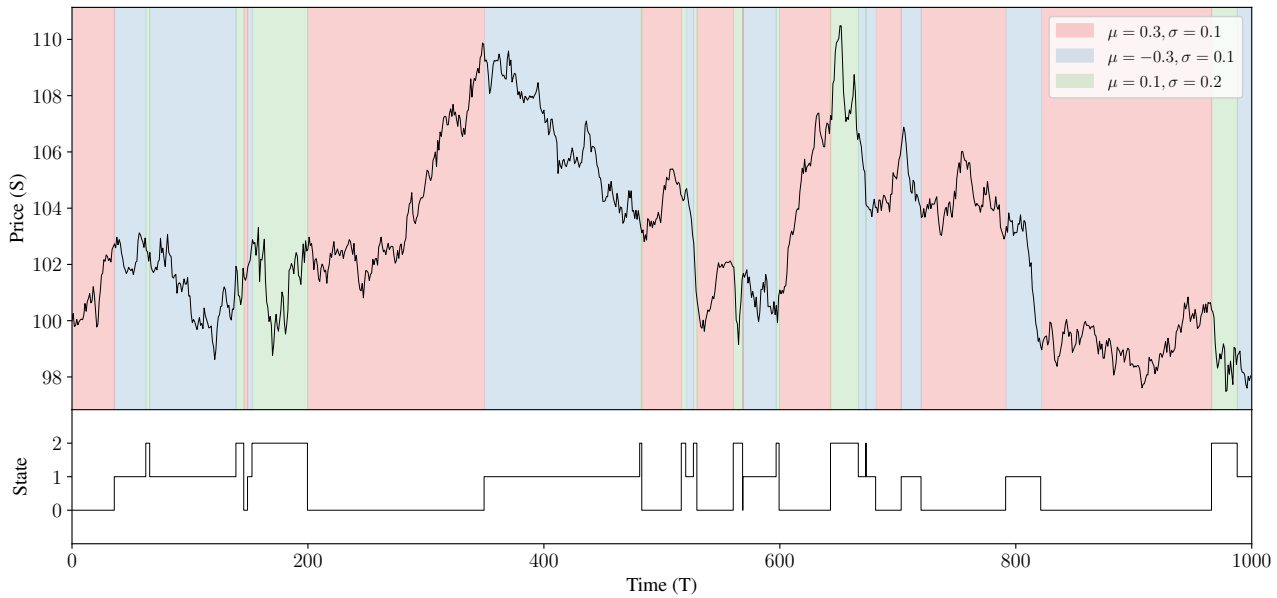


Figure 2: An example path of a regime-switching geometric Brownian motion

u. Therefore, given u, the exponential random variable t_1 is given by:

$$t_1 = \frac{-\ln(u/\lambda_1)}{\lambda_1}. \tag{39}$$

Let T_1 be the occupation time of the Markov chain in state 1 during the interval $[0, T]$ and T_2 be the occupation of chain in state 2. We know that $T_1 = t_1 + (t_3 - t_2)$ and $T_2 = (t_2 - t_1) + (T - t_3)$. Therefore, we have

$$\begin{aligned}\int_0^T \sigma_s \, dW_s &= \int_0^{t_1} \sigma_s \, dW_s + \int_{t_1}^{t_2} \sigma_s \, dW_s + \int_{t_2}^{t_3} \sigma_s \, dW_s + \int_{t_3}^{T} \sigma_s \, dW_s \\ &= \sigma_1(W_{t_1} - W_0) + \sigma_1(W_{t_2} - W_{t_1}) + \sigma_1(W_{t_3} - W_{t_2}) + \sigma_1(W_T - W_{t_3}).\end{aligned}$$

Using the independent increments property of Brownian motion and rules for manipulating normal distributions, we have

$$\sigma_1(W_{t_1} - W_0) + \sigma_1(W_{t_3} - W_{t_2}) \sim \mathcal{N}\left(0, \sigma_1^2(t_1 + (t_3 - t_2))\right) = \mathcal{N}\left(0, \sigma_1^2 T_1\right). \tag{40}$$

Using the above implies that

$$\begin{aligned}\ln(S_t) &\overset{\mathrm{d}}{=} \ln(S_0) + \int_0^{T_1} (\mu_1 - \frac{1}{2}\sigma_1^2)\, ds + \int_{T_1}^{T} (\mu_2 - \frac{1}{2}\sigma_2^2)\, ds + \int_0^{T_1} \sigma_1 \, dW_s + + \int_{T_1}^{T} \sigma_2 \, dW_s \\ &\overset{\mathrm{d}}{=} \ln(S_0) + (\mu_1 - \frac{1}{2}\sigma_1^2)T_1 + (\mu_2 - \frac{1}{2}\sigma_2^2)(T - T_1) + \sigma_1 W(T_1) + \sigma_2(W(T) - W(T_1)) \\ &\sim \mathcal{N}\left(\ln(S_0) + (\mu_1 - \frac{1}{2}\sigma_1^2)T_1 + (\mu_2 - \frac{1}{2}\sigma_2^2)(T - T_1), \sigma_1^2 T_1 + \sigma_2^2(T - T_1)\right).\end{aligned}$$

Therefore, we can take each sample of the log-terminal stock price using only one pseudo-random sample from the standard normal distribution. Let N be the number of runs. For $n = 1, \ldots N$,

1. Obtain the n-th sample path of η_t, for $t \in [0, T]$,
2. Calculate the occupation time in state 1, T_1 and state 2, $T - T_1$,
3. Simulate the log-terminal stock price by taking a sample from the distribution,
4. Calculate the payoff of the call option: $C_n(K) = e^{-rT} \max(S_T - K, 0)$,
5. Calculate an unbiased sample mean $\hat{C}(K)$ of the price of the option using: $\hat{C}(K) = \frac{1}{N} \sum_{n=1}^{N} C_n(K)$.

C Obtaining Market Data

The regime-switching European call option formula is ready to be tested on real market data. We note that the choice of market and type of instrument is important due to some of the simplifying assumptions made in the Black-Scholes universe. We have assumed that when dealing with the risky asset (the underlying), cash or the option itself, we are not subject to any commission, transfer fees, option exercise or assignment fees. There is also the assumption that the bid-ask spread in the market is negligible. With this in mind, we use the midpoint between the bid and the ask to represent the market price. We also assumed perfect liquidity, whereby it is possible to buy or sell any quantity of any of our assets at any time. Therefore, we choose to calibrate the model against the quoted market prices of call options on the SPY ETF. The SPY index is simply the numerical value that represents the level of the S&P 500 index. The option prices are obtained using *yfinance*, a Python module for downloading market data from Yahoo! Finance, as of Friday 15th September 2023. The price of the underlying is set at $443.37, the closing price on the same day.

D Determining the Risk-free Rate

In the context of fixed income securities like bonds, the benchmark yield usually refers to the yield on a standard security (often a government bond like a U.S. Treasury) against which other securities are measured. The benchmark yield is the bond yield that is quoted by the market on the sovereign/government bonds trading in that market. It is termed the benchmark owing to the fact that this is for a sovereign bond with specific maturity (generally 10 years). This yield is like a proxy yield that is used by investors to get a sense of the interest rates in the economy. Just like 10Y yield, there are sovereign securities with other maturities that are traded in the market.

The Nelson-Siegel model is a popular mathematical framework used for modeling the shape of the yield curve in the fixed-income market [24]. The model expresses the zero-coupon yield $y(t)$ at time t as a function of four parameters β_0, β_1, β_2 and τ:

$$y(t) = \beta_0 + \beta_1 \left(\frac{1 - e^{-t/\tau}}{t/\tau} \right) + \beta_2 \left(\frac{1 - e^{-t/\tau}}{t/\tau} - e^{-t/\tau} \right) \tag{41}$$

where β_0 is the long-term factor, β_1 is the short-term factor, β_2 is the curvature factor and τ is the decay factor. There is however one drawback of the Nelson-Siegel model: it is unable to handle a change in slope and curvature along the tenors. An extension to the aforementioned approach, the Nelson-Siegel-Svensson model defines the the zero-coupon yield $y(t)$ at time t as a function of six parameters β_0, β_1, β_2, β_3, τ_1 and τ_2:

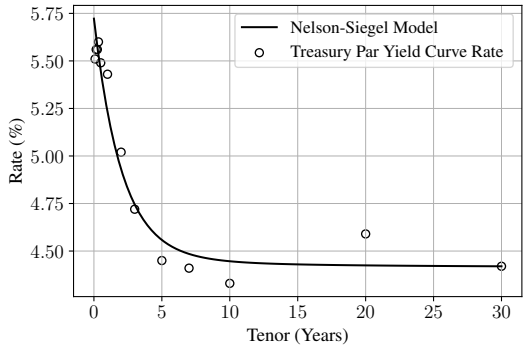
$$y(t) = \beta_0 + \beta_1 \left(\frac{1 - e^{-t/\tau_1}}{t/\tau_1} \right) + \beta_2 \left(\frac{1 - e^{-t/\tau_1}}{t/\tau_1} - e^{-t/\tau_1} \right) + \beta_3 \left(\frac{1 - e^{-t/\tau_2}}{t/\tau_2} - e^{-t/\tau_2} \right) \tag{42}$$

where β_3 is the second-curvature factor and τ_2 is the second decay factor. In the interest of efficiency, the Nelson-Siegel and Nelson-Siegel-Svensson models are used via the *nelson-siegel-svensson* Python module. Although we only explore the calibration of options with maturities of less than three years, these curves would be necessary when endeavouring to price Long-Term Equity Anticipation Securities (LEAPS) for example. Visual examples of these fitted curves can be found in Figure 3.

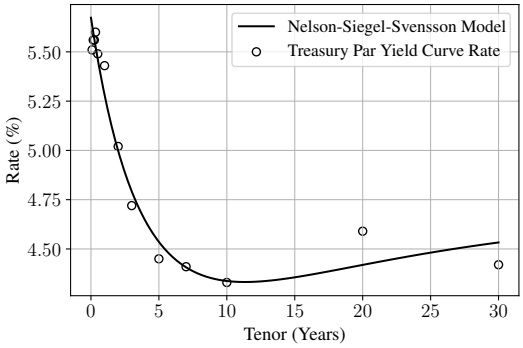
E Calibration of the Model

With a stable implementation of closed-form solution of the regime-switching model, we are able to calibrate the model to the prices of real SPY call options. Calibration is carried out using least squares minimisation, adapting the method proposed by Mikhailov and Nögel [25]. Suppose we have a vector Θ containing the four parameters we wish to calibrate: $\Theta = (\lambda_1, \lambda_2, \sigma_1, \sigma_2)$. Let $T = (T_1, T_2, \ldots, T_M)$ be M times to maturity, and let $K = (K_1, K_2, \ldots, K_N)$ be N strike prices. The aim of the calibration is to find some solution vector $\hat{\Theta}$ such that that the squared error is minimised:

$$\hat{\Theta} = \arg\min_{\Theta} \sum_{(T,K) \in T \times K} [C(K,T) - C_{\text{RS}}(K,T)]^2 + \text{Penalty}(\Theta, \Theta_0) \tag{43}$$



(a) Nelson-Siegel model: $\beta_0 = 0.0441$, $\beta_1 = 0.0131$, $\beta_2 = -0.0117$, $\tau = 2.0$



(b) Nelson-Siegel-Svensson model: $\beta_0 = 0.0484$, $\beta_1 = 0.00832$, $\beta_2 = -0.00147$, $\beta_3 = -0.0216$, $\tau_1 = 2.0$, $\tau_2 = 5.0$

Figure 3: Calibrated Treasury Par Yield Curve

where $C(K,T)$ denotes the market price for a call with strike K and time to maturity T. $C_{\text{RS}}(K,T)$ is the price calculated by the regime-switching model. Θ_0 is the vector containing the four starting values of our parameters and is used to calculate the penalty function, which is used to give the calibration additional stability; we use $\text{Penalty}(\Theta, \Theta_0) = ||\Theta - \Theta_0||^2$.

F Calculating the Implied Volatility

The implied volatility is calculated using the aforementioned Newton-Raphson scheme. Due to the repeated use of this implied volatility calculation during calibration, it is prudent to endeavour to reduce the time it takes. The Newton-Raphson method converges more quickly when the initial guess, σ_0, is close to the root σ. Various approximations for the volatility are proposed throughout the literature and serve as a useful initial guess for our numerical method. One commonly used closed-form estimate was proposed by Brenner and Subrahmanyam [26]. They approximate the volatility as

$$\sigma \approx \sqrt{\frac{2\pi}{T}} \cdot \frac{C}{S} \tag{44}$$

where C is the price of the European call option, T is the time to maturity, and S is the price of the underlying. Their approximation is derived using the fact that, for an at-the-money option, $S = Ke^{-rT}$. Therefore, the approximation does break down as K moves further away from S. For example, with the parameters $C = 22.51$, $S = 100$, $T = 1$, our initial guess $\sigma_0 \approx 0.564$. The implementation used has a tolerance $\epsilon = 0.001$ and a maximum number of iterations of 200. Figure 4 shows the convergence of the method within two iterations.

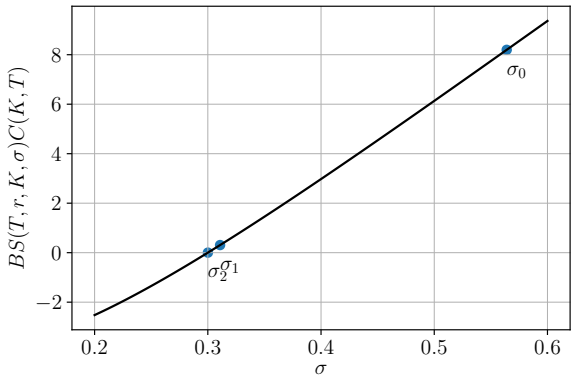


Figure 4: The Newton-Raphson method for calculating implied volatility σ with call price $C = 22.51$, stock price $S = 100$ and time to expiry $T = 1$

IV RESULTS AND EVALUATION

A *Verification of the Continuous-Time Markov Chain*

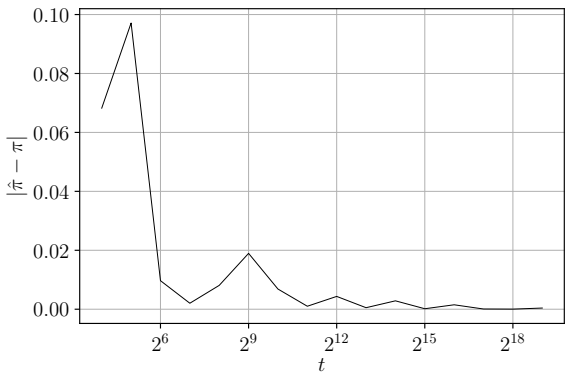
A significant component of the model is the underlying continuous-time Markov chain. Therefore, it is important to be able to quantitatively measure the implementation of the chain and verify that it is indeed correct. This verification can be achieved by simulating the Markov chain for a long period of time and checking that the amount of time spent in each state broadly matches the expected stationary distribution π of the Markov chain. For the sake of simplicity, and to use the previous derivation of the stationary distribution for a two-state continuous-time Markov chain, consider again a two-state chain with generator matrix with real values assigned to λ_1 and λ_2:

$$Q = \begin{bmatrix} -1 & 1 \\ 3 & -3 \end{bmatrix}. \tag{45}$$

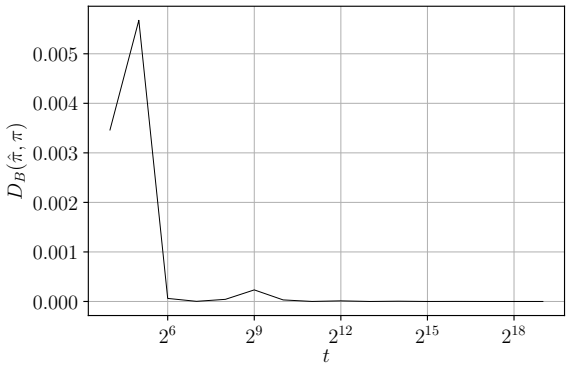
Using equation, we expect the stationary distribution of the simulated chain, as the length of the simulation increases, to approach $\pi = (\frac{3}{4}, \frac{1}{4})$ i.e. the chain spends 75% of the time in state 1 and 25% of the time in state 2. The distribution that emerges from simulating a continuous-time Markov chain with generator matrix Q is calculated by calculating the ratio of time spent in each state, and is termed the emergent stationary distribution. This is done for increasing time t to demonstrate the convergence of the emergent stationary distribution $\hat{\pi}$ towards the expected stationary distribution π. The Bhattacharyya distance is used to calculate the distance between the two distributions so that we can establish whether they get closer together as t gets larger. The Bhattacharyya distance between two probability distributions P and R on the same domain χ is defined as

$$D_B(P, R) = -\ln \sum_{x \in \chi} \sqrt{P(x)R(x)}. \tag{46}$$

Figure 5 clearly shows that both the absolute error and the Bhattacharyya distance approach zero as $t \to \infty$ for the two state Markov chain. We can therefore be confident that our implementation is correct. The implemen-



(a) Absolute error, $|\hat{\pi} - \pi|$, between the emergent and expected stationary distributions against the length of the simulation t



(b) Bhattacharyya distance, $D_B(\hat{\pi}, \pi)$, between the emergent and expected stationary distributions against the length of the simulation t

Figure 5: Convergence of the 2-state continuous-time Markov chain

tation also works for any number of states. Consider the generator matrix Q for a three-state continuous-time Markov chain:

$$Q = \begin{bmatrix} -6 & 3 & 3 \\ 4 & -12 & 8 \\ 15 & 3 & -18 \end{bmatrix}. \tag{47}$$

As discussed previously, the stationary distribution of a given generator matrix Q can be determined by solving the linear equation:

$$\pi Q = 0, \quad \sum_{i \in S} \pi_i = 1. \tag{48}$$

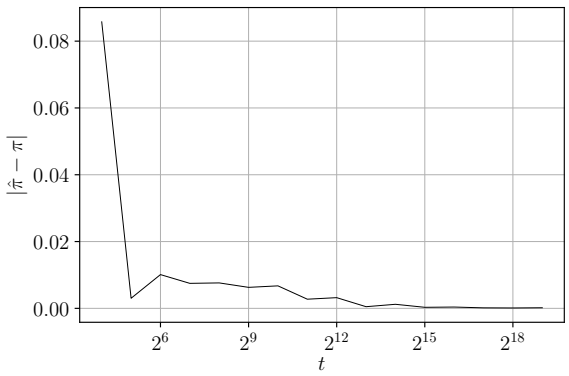
This is achieved first by augmenting Q to get M in order to include the extra constraint and defining b as

$$M = \begin{pmatrix} \hat{Q}^T \\ 1...1 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \tag{49}$$

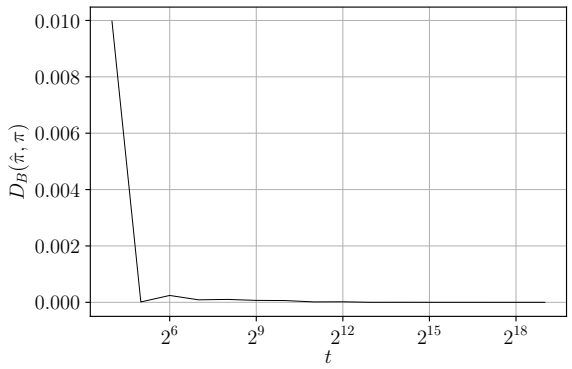
where $\hat{Q}$ is the Q matrix with the last column removed. The intuition behind this augmentation is that, by removing the bottom row and replacing it with $\mathbb{1}_n = (1, ..., 1)$, we incorporate the constraint, while retaining a system of only three equations with three unknowns. After this augmentation, the stationary distribution of the Markov chain π is the solution to

$$M\pi = b \tag{50}$$

which can be solved computationally as $\pi = (0.6095..., 0.2, 0.1904...)$. Having determined the expected stationary distribution for our three-state continuous time Markov chain, we can plot its absolute error and Bhattacharyya distance against the length of the simulation in Figure 6. The three-state simulated chain's stationary



(a) Absolute error, $|\hat{\pi} - \pi|$, between the emergent and expected stationary distributions against the length of the simulation t



(b) Bhattacharyya distance, $D_B(\hat{\pi}, \pi)$, between the emergent and expected stationary distributions against the length of the simulation t

Figure 6: Convergence of the 3-state continuous-time Markov chain

distribution does seem to approach the expected distribution, verifying that our implementation generates valid outputs. We can be confident that the trajectories our implementation generates will, on average, be in each state for the correct amount of time.

B Comparing Analytical and Monte-Carlo Prices

A common technique for verifying the analytical solution for our European call option price is to compare it with our Monte-Carlo price and evaluate the error. For our experiments, we use the following parameters: $S_0 = 100$, $K = 90$, $\lambda_1 = \lambda_2 = 1$, $\mu_1 = \mu_2 = 0.1$, $r = 0.1$, $\sigma_1 = 0.2$, $\sigma_2 = 0.3$ and $N = 10^6$. Table 1 shows the numerical results of our Monte-Carlo experiments when the initial state of the continuous-time Markov chain is 1. Table 2 shows the numerical results for the same maturities but with an initial state of 2. It is clear

Time to maturity	Analytical price	Monte-Carlo price	Error	95% Confidence Interval
0.1	10.993	11.002	0.009	(10.963 - 11.041)
0.2	12.165	12.144	-0.021	(12.089 - 12.199)
0.5	15.614	15.625	0.011	(15.534 - 15.717)
1.0	20.722	20.768	0.046	(20.624 - 20.912)
2.0	29.288	29.286	-0.002	(29.041 - 29.530)
3.0	36.477	36.623	0.147	(36.273 - 36.974)

Table 1: Comparison of Analytical and Monte-Carlo Prices when $\eta_0 = 1$

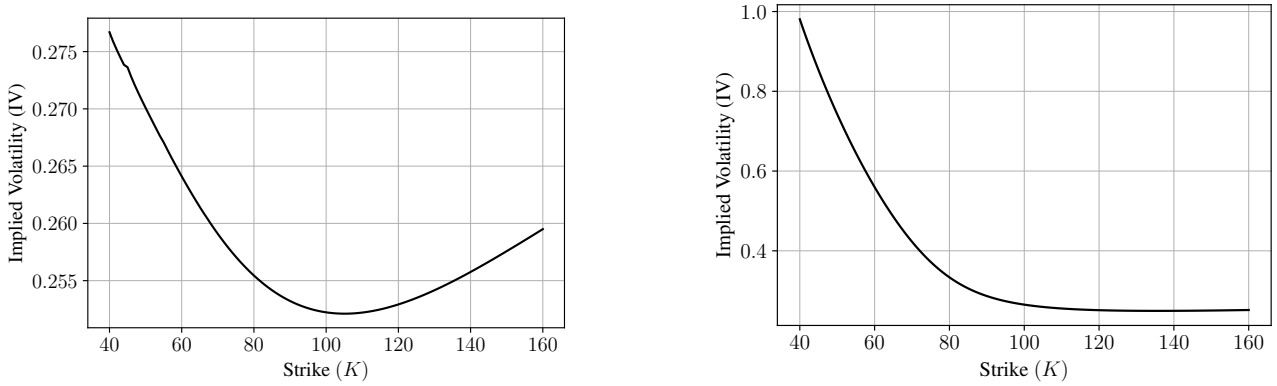
Time to maturity	Analytical price	Monte-Carlo price	Error	95% Confidence Interval
0.1	11.361	11.360	-0.001	(11.306 - 11.414)
0.2	12.889	12.935	0.045	(12.861 - 13.008)
0.5	16.718	16.799	0.081	(16.685 - 16.914)
1.0	21.812	21.891	0.079	(21.723 - 22.059)
2.0	30.085	30.002	-0.083	(29.737 - 30.267)
3.0	37.062	37.048	-0.014	(36.680 - 37.416)

Table 2: Comparison of Analytical and Monte-Carlo Prices when $\eta_0 = 2$

from Tables 1 and 2 that the Monte-Carlo simulations do indeed converge to the analytical prices. The error does not increase as the time to maturity, T, increases, indicating the accuracy of the Monte-Carlo simulation for both near-to- and far-from-expiry options alike.

C Volatility Smile and Term Structure

Having confirmed that our analytical and experimental prices for European call options are in agreement, we can proceed with illustrating the volatility smile and volatility term structures implied by the model. As shown in Figure 7, the volatility implied by the prices under the regime-switching Black-Scholes varies by strike. This



(a) $\lambda_1 = 1, \lambda_2 = 1$

(b) $\lambda_1 = 1, \lambda_2 = 2$

Figure 7: Volatility smile/skew ($T = 1$, $S_0 = 100$, $\sigma_1 = 0.2$, $\sigma_2 = 0.3$, $\lambda_1 = \lambda_2 = 1$, $r = 0.05$)

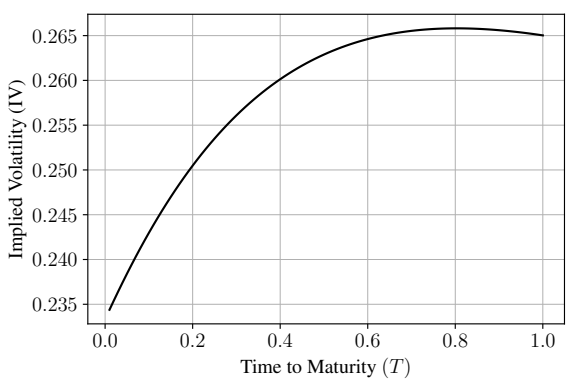
non-constant volatility mimics the implied volatility found in real markets. The volatility smile in Figure 7a can often be seen in foreign exchange options or equity index options. The volatility skew in Figure 7b is more common in equity option markets.

Figure 8 shows the implied volatility term structure of the model with the accompanying parameters as input. Hollstein et al. [27] discuss various assets and the general shape of their implied volatility term structure. In Figure 8a, the volatility increases as the time to maturity increases; this is often found in equity markets. Options that have commodities, in particular non-metal commodities, such as crude oil and natural gas, as their underlying tend to have a decreasing volatility term structure similar to that of Figure 8b.

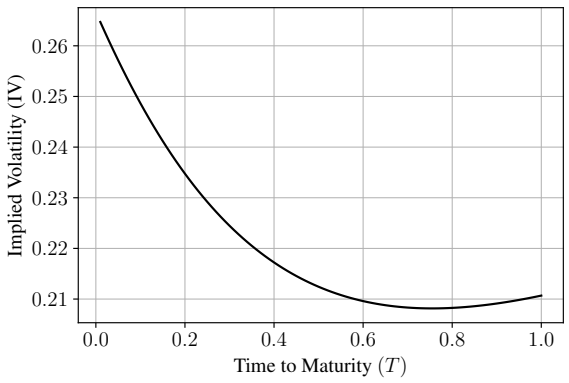
D Price Calibration

The regime-switching model is calibrated using the calibration scheme described previously. As established, the model is calibrated against quoted market prices of European call options on the SPY ETF as of Friday 15th September 2023.

The parameter vector $\Theta_0 = (\lambda_1, \lambda_2, \sigma_1, \sigma_2)$ was initialised after examining similar experiments done by He and Zhu [28]. These starting parameters and their lower and bounds are detailed in Table 3. The lower bounds for σ_1 and σ_2 are self explanatory: we cannot have negative volatility. The lower bounds for λ_1 and λ_2 are set to be greater than zero to avoid an absorbing state situation. The simplifying case when $\lambda_2 = 0$ was investigated by Yao et al. [29]. Least squares minimisation is carried out using the *scipy.minimize* function.



(a) $\lambda_1 = 1, \lambda_2 = 2$



(b) $\lambda_1 = 2, \lambda_2 = 1$

Figure 8: Volatility term structure ($K = 100$, $S_0 = 100$, $\sigma_1 = 0.2$, $\sigma_2 = 0.3$, $r = 0.05$)

Parameter	Initial Value (Θ_0)	Lower Bound	Upper Bound	Final Value ($\hat{\Theta}$)
σ_1	0.1	0.01	1	0.0776
σ_2	0.2	0.01	1	0.255
λ_1	0.8	0.01	5	0.547
λ_2	2	0.01	5	1.92

Table 3: Parameters with Initial Values and Bounds

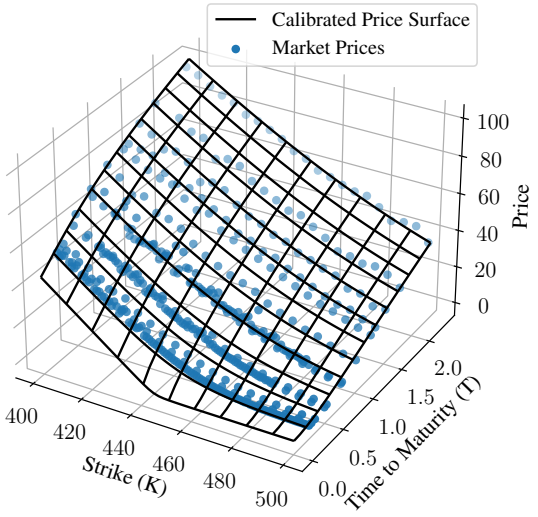
Various optimisation algorithms are available to carry out bound-constrained minimisation; we decided to use L-BFGS-B [30]. The optimised parameter vector $\hat{\Theta}$ constitutes a two-state model where one state is a high volatility environment, and the other is a low volatility environment. Using λ_1 and λ_2, we can, as before, calculate the expected stationary distribution of the model:

$$\pi_1 = \frac{\lambda_2}{\lambda_1 + \lambda_2} = 0.778, \qquad \pi_2 = \frac{\lambda_1}{\lambda_1 + \lambda_2} = 0.222. \tag{51}$$

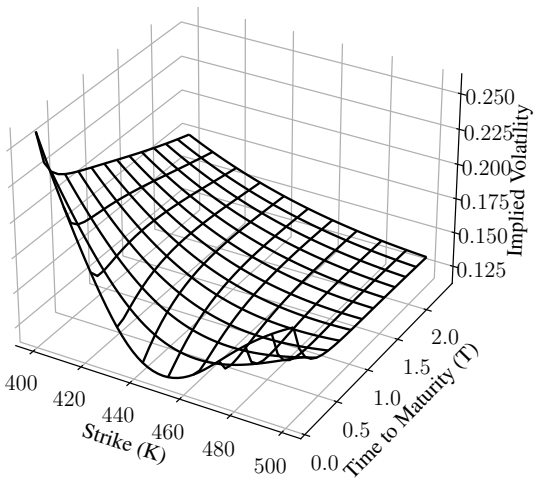
We can interpret our parameters as describing a two-state model that assumes that 77.8% of the time before expiry T, the underlying has a volatility of 0.0776, and 22.2% of the time, the underlying has a volatility of 0.255. Figure 9a shows the market prices and the fitted price surface. Figure 9b shows the volatility surface implied by the prices calculated in Figure 9a. The shape of this volatility surface is encouraging. It demonstrates the non-constant volatility implied by the model rather than the constant volatility that a vanilla Black-Scholes model assumes. We can see that the model has fit the market prices well in most cases.

The main drawback of the model however is that it fails to accurately price European call options that are near-to-expiry and deeply out-of-the-money (when the strike price K is significantly higher than the current underlying price S). The price of an option can be thought of as the sum of its intrinsic and extrinsic (or time) values. Out-of-the-money calls have no intrinsic value (as $K > S$, so the payoff is 0) and only possess extrinsic value. The extrinsic value of an option rises as volatility rises, so we can conclude that our model is overestimating the volatility for these options. This pricing error can be seen in Figure 10 where the percentage error is plotted against the strike and time to maturity.

However, for call options that are deeply in-the-money, the error produced by the model is significantly lower. This is because they have intrinsic value which is the difference between the price of the underlying and the strike: $S - K$. For example, a call option with strike $K = 420$ has an intrinsic value of $\$443.37 - \$420 = \$13.37$, a large proportion of its price. The value of these deeply in-the-money options are weighted more heavily towards their intrinsic value than their extrinsic value. Therefore, as these options are priced more accurately, it is likely that the model unfortunately struggles to calculate extrinsic value when the time to maturity is small. On the contrary, if the time to maturity is larger than $T = 1$, the model does successfully price the option irrespective of its strike K. With these results in mind, we hypothesise that there may be an error with the implementation of the numerical integration when T is small.



(a) Calibrated price surface with market prices



(b) Calibrated volatility surface

Figure 9: Calibration results

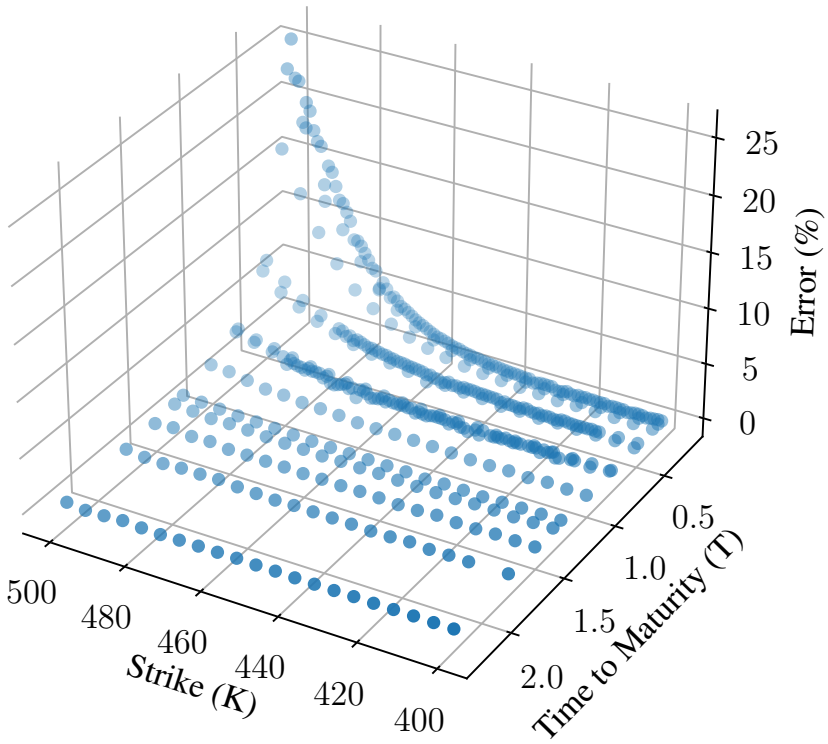


Figure 10: Price error

V CONCLUSION

This paper investigates the use of the regime-switching Black-Scholes model to predict the price of European call options on the SPY ETF. We explore the various shortcomings of the vanilla Black-Scholes model, namely its assumption of constant volatility over time. This motivates us to explore various ways of incorporating variable volatility into the model, including stochastic volatility models and regime-switching models. We then give basic derivations of the analytic solution for the price of European call option when the supporting continuous-time Markov chain has two states. We validate our continuous-time Markov chain but plotting the Bhattacharyya distance between the expected stationary distribution, given λ_1 and λ_2, and the emergent stationary distribution. We see that the error goes to zero as you run the Markov chain for an increasing length of time.

As is common in the field, we devise a Monte Carlo scheme to calculate numerically the prices and implement this in Python. We have found that the analytical and Monte Carlo prices are very close; all of our results, whether starting in state 1 or state 2, have an error of less than 0.5%. These errors would reduce as the number of experiments, N, was increased beyond 10^6. In the interest of exploring the use of such models in industry, we calibrate the regime-switching model to the market prices of SPY call options as of Friday 15th September 2023. The model fits the data with varying success depending on the strike K and the time to maturity T. Industry applications of this type of option pricing models are countless. They can be used by market-makers to calculate the fair price of an option in the market so that they can publish a bid and an ask price. Market makers capture the bid-ask spread while remaining risk neutral by using option pricing models to hedge their portfolios. Risk takers, such as hedge funds or proprietary trading firms, may use the same models to identify mispriced options in the market and carry out arbitrage [31].

The paper overall addresses all the main objectives identified at the outset. The success of the work hinges mainly on the calibration results. While the calibrated model prices in-the-money and longer maturity options well, it is less accurate for out-of-the-money options that are very close to expiry. Further experiments should be carried out to investigate the reason for this.

There are several directions that could be investigated to further this work. The quality of the price data used for calibration has a large impact on quality of the fit; obtaining data from a paid provider would hopefully mitigate stale or incorrect data. The minimisation can also be improved: ordinary least squares is currently used for fitting the model to the market prices. Weighted least squares could be used to improve the calibration; Homescu [32] provides two options. The inverse square of the bid-ask spread can be used to give greater weight to options with better liquidity. The other approach is to weight the options by the inverse of its Vega. The author suggests that it might be better to minimise the differences of implied volatilities rather than prices in order to avoid the minimal weighting towards options that are far-from-the-money (due to increased bid-ask spreads). The author seeks to reduce computational complexity by approximating this as minimising the square differences of option prices weighted by the Black-Scholes Vega evaluated at the implied volatility of the market option price. Further work could also be done to quantitatively compare the regime-switching model to a stochastic volatility model such as Heston's using the same price data.

The model's application to options with more varied underlyings would also provide value. Whilst this work explores pricing European call options on a single index-tracking ETF, further work could explore the pricing of options on indices in emerging markets, where the expected rate of return μ and, more importantly, the volatility σ can be significantly higher. Options on commodity future contracts, foreign exchange rates and interests could also be explored. Finally, although not necessary in this work due to the availability and relative ease of calculating the analytic solutions, Monte Carlo methods could be used to price more exotic options that are normally traded over-the-counter, such as Asian, Bermudan, lookback and binary options [33].

References

[1] G. W. Schwert, "Why does stock market volatility change over time?" *The Journal of Finance*, vol. 44, no. 5, pp. 1115–1153, 1989.

[2] R. F. Engle and A. J. Patton, *Forecasting Volatility in the Financial Markets*, Third Edition. Oxford, 2007.

[3] M. J. Lorig and R. Sircar, "Stochastic volatility: Modeling and asymptotic approaches to option pricing and portfolio selection," 2016.

[4] J. D. Hamilton, "A new approach to the economic analysis of nonstationary time series and the business cycle," *Econometrica*, vol. 57, no. 2, pp. 357–384, 1989.

[5] J.-P. Fouque, G. Papanicolaou, and K. Sircar, "Mean-reverting stochastic volatility," *International Journal of Theoretical and Applied Finance*, vol. 3, Feb. 1999.

[6] J.-P. Fouque and G. C. Papanicolaou, "Stochastic volatility correction to Black-Scholes," 2000.

[7] F. Black and M. Scholes, "The pricing of options and corporate liabilities," *Journal of Political Economy*, vol. 81, no. 3, pp. 637–654, 1973.

[8] R. C. Merton, "Theory of rational option pricing," *The Bell Journal of Economics and Management Science*, vol. 4, no. 1, pp. 141–183, 1973.

[9] C. Eom, T. Kaizoji, and E. Scalas, "Fat tails in financial return distributions revisited: Evidence from the Korean stock market," *Physica A: Statistical Mechanics and its Applications*, vol. 526, p. 121 055, 2019.

[10] B. Eraker, "Do stock prices and volatility jump? Reconciling evidence from spot and option prices," *The Journal of Finance*, vol. 59, no. 3, pp. 1367–1403, 2004.

[11] R. C. Merton, "Option pricing when underlying stock returns are discontinuous," *Journal of Financial Economics*, vol. 3, no. 1, pp. 125–144, 1976.

[12] J. Hull, *Options, futures, and other derivatives*. 1997.

[13] M. Haugh, *The Black-Scholes model*, 2016.

[14] S. Natenberg, *Option Volatility and Pricing: Advanced Trading Strategies and Techniques*, 2nd ed. 2014.

[15] D. R. Chambers and S. K. Nawalkha, "An improved approach to computing implied volatility," *SSRN Electronic Journal*, 2009.

[16] S. L. Heston, "A closed-form solution for options with stochastic volatility with applications to bond and currency options," *The Review of Financial Studies*, vol. 6, no. 2, pp. 327–343, 1993.

[17] D. Anderson and T. Kurtz, "Continuous time Markov chain models for chemical reaction networks," in Apr. 2011, pp. 3–42.

[18] C.-D. Fuh, K. W. R. Ho, I. Hu, and R.-H. Wang, "Option pricing with markov switching," *Journal of Data Science*, vol. 10, no. 3, pp. 483–509, 2022.

[19] J. D. Hamilton, "Rational-expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates," *Journal of Economic Dynamics and Control*, vol. 12, no. 2, pp. 385–423, 1988.

[20] X. Guo, "Information and option pricings," *Quantitative Finance*, vol. 1, no. 1, pp. 38–44, 2001.

[21] V. Naik, "Option valuation and hedging strategies with jumps in the volatility of asset returns," *The Journal of Finance*, vol. 48, no. 5, pp. 1969–1984, 1993.

[22] E. Sohrabi, "Option valuation under a regime-switching model using the fast fourier transform," 2018.

[23] "Continuous-time stochastic calculus," in *Mathematics of Financial Markets*. New York, NY: Springer New York, 2005, pp. 131–165.

[24] J. Annaert, A. G. P. Claes, M. J. K. de Ceuster, and H. Zhang, "Estimating the yield curve using the Nelson-Siegel model: A ridge regression approach," *International Review of Economics & Finance*, 2012.

[25] S. S. Mikhailov and U. Nögel, "Heston's stochastic volatility model implementation, calibration and some extensions," 2003.

[26] M. Brenner and M. Subrahmanyam, "A simple formula to compute the implied standard deviation," *Financial Analysts Journal*, vol. 44, pp. 80–83, Sep. 1988.

[27] F. Hollstein, M. Prokopczuk, and C. Würsig, "Volatility term structures in commodity markets," *Journal of Futures Markets*, vol. 40, no. 4, pp. 527–555, 2020.

[28] X.-J. He and S.-P. Zhu, "How should a local regime-switching model be calibrated?" *Journal of Economic Dynamics and Control*, vol. 78, pp. 149–163, 2017.

[29] D. Yao, Q. Zhang, and X. Zhou, "A regime-switching model for european options," in *Stochastic Processes, Optimization, and Control Theory: Applications in Financial Engineering, Queueing Networks, and Manufacturing Systems*. Sep. 2006, pp. 281–300.

[30] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A limited memory algorithm for bound constrained optimization," *SIAM Journal on Scientific Computing*, vol. 16, no. 5, pp. 1190–1208, 1995.

[31] M. Boffetti, *Option volatility & arbitrage opportunities*, 2016.

[32] C. Homescu, *Implied volatility surface: Construction methodologies and characteristics*, 2011.

[33] Y. An and W. Suo, "An empirical comparison of option-pricing models in hedging exotic options," *Financial Management*, vol. 38, no. 4, pp. 889–914, 2009.