# Beyond Black-Scholes: Moving to Stochastic Volatility

Nikhil A Shah

February 21, 2023

## Contents

# 1 Introduction

For each model that is presented, we will begin by analysing the stochastic differential equation that governs the movement of the underlying asset price. We will then move towards an explicit solution (which in each of these cases will follow some kind of exponential Lèvy process), before working toward understanding the price of contingent claims based upon these underlying movements. We will only focus on the pricing of European options since we have closed-form solutions that we can analyse effectively.

In Section 2, we delve into some of the details of the Black-Scholes model and in particular assess the assumption that volatility is treated as a constant parameter by looking at real world data. In Sections 3 and 4 we allow extensions of the Black-Scholes model by first incorporating stochastic volatility as in the Heston model, before adding jumps into the path of the underlying as in the Merton model.

The code that accompanies this paper and generates the outputs can be found at `https://github.com/shahnikhil12/Simulation-of-Stochastic-Processes`

# 2 Black-Scholes Model

## 2.1 Set-Up

Let $(S_t)_{t\geqslant 0}$ denote the price of an underlying stock that is initialised with starting data given by $S_0$. For simplicity we will assume that $S_t$ is one-dimensional, ie we are only modelling the movement of one underlying asset. The exposition below extends simply in the case of multiple assets, mutatis mutandis. Let $\mu \in \mathbb{R}$ denote the percentage drift of the underlying, and let $\sigma \in \mathbb{R}_{>0}$ be the percentage volatility. For the time begin we will assume that these are constants. Further, let $(W_t)_{t\geqslant 0}$ be a scalar Brownian Motion. Then the underlying asset price evolves under the following stochastic differential equation:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{1}$$

Whilst the details of the solution to this stochastic differential equation are not of prime importance, they will be discussed in Appendix A. It will be sufficient for the rest of the analysis to quote that the solution of Equation 1 is given by

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

From here, we can now move on to discuss option pricing in this model. Broadly speaking, there are two ways of moving toward deriving the famous Black-Scholes formula. One method is to make use of facts about a quantity known as the state price density and then apply the Cameron-Martin-Girsanov theorem to switch to a risk neutral measure and compute a specific integral. Whilst this calculation is fairly simple, it requires a large number of auxiliary definitions and results and is thus reserved to Appendix B. Hence we shall use a second method that relies on a duality between the Feynman-Kac formula and the Feynman-Kac partial differential equation.

### 2.1.1 Deriving Black-Scholes Formula

For simplicity, we shall work with a version of Equation (1) where we let $\mu = r$ where $r$ is the understood as the interest rate. Further, we will be interested in a function

$$V : [0,T] \times \mathbb{R} \to [0,\infty) \tag{3}$$
$$(t, S_t) \mapsto V(t, S_t) \tag{4}$$

where we can understand $V$ as a 'value' of the option. By applying Itô's Lemma, we can begin analysing the differential of $V$:

$$dV(t,S_t) = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS_t + \frac{1}{2}\frac{\partial^2 V}{\partial t^2}\langle dt, dt\rangle + \frac{\partial^2 V}{\partial t \partial S}\langle dt, dS_t\rangle + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\langle dS_t, dS_t\rangle \tag{5}$$

where the angular brackets denote quadratic covariation. We start by stating the quadratic covariation identities for $dt$ and $dW_t$ which can be derived using elementary stochastic calculus and measure theory

$$\langle dt, dt\rangle = 0 \tag{6}$$
$$\langle dt, dW_t\rangle = 0 \tag{7}$$
$$\langle dW_t, dW_t\rangle = dt \tag{8}$$

Hence we can derive the covariation identities as needed in Equation (5):

$$\langle dt, dS_t\rangle = rS_t\langle dt, dt\rangle + \sigma S_t\langle dt, dW_t\rangle = 0 \tag{9}$$
$$\langle dS_t, dS_t\rangle = r^2S_t^2\langle dt, dt\rangle + 2r\sigma S_t\langle dt, dW_t\rangle + \sigma^2S_t^2\langle dW_t, dW_t\rangle = \sigma^2S_t^2 dt \tag{10}$$

Hence, we can simplify Equation (5) by using Equation (1) and Equations (9)-(10):

$$dV(t,S_t) = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}(rS_t dt + \sigma S_t dW_t) + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\sigma^2 S_t^2 dt \tag{11}$$
$$= \left(\frac{\partial V}{\partial t} + rS_t\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2S_t^2\frac{\partial^2 V}{\partial S^2}\right)dt + \sigma S_t\frac{\partial V}{\partial S}dW_t \tag{12}$$

We can then apply the product rule to the discounted process $(e^{-rt}V(t,S_t))_{t\geqslant 0}$ to finds its differential:

$$d\left\{e^{-rt}V(t,S_t)\right\} = e^{-rt}dV(t,S_t) - re^{-rt}V(t,S_t)dt \tag{13}$$
$$= e^{-rt}\left(\frac{\partial V}{\partial t} + rS_t\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2S_t^2\frac{\partial^2 V}{\partial S^2}\right)dt + e^{-rt}\sigma S_t\frac{\partial V}{\partial S}dW_t - re^{-rt}V(t,S_t)dt \tag{14}$$
$$= \left(\frac{\partial V}{\partial t} + rS_t\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2S_t^2\frac{\partial^2 V}{\partial S^2} - rV\right)e^{-rt}dt + e^{-rt}\sigma S_t\frac{\partial V}{\partial S}dW_t \tag{15}$$

Thus, by requiring that the discounted process $(e^{-rt}V(t, S_t))_{t\geqslant 0}$ is a local martingale, we require that the drift term must vanish. This gives us the celebrated Black-Scholes PDE:

$$\frac{\partial V}{\partial t} + rS_t\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 V}{\partial S^2} = rV \tag{16}$$

Equation (16) is the Feynman-Kac PDE for the Black-Scholes model. In order to to find the value of the call option, we rely on the following theorem (the theorem and its proof will work in a more general setting where $S_t \in \mathbb{R}^n$)

**Theorem 2.1** (Feynman-Kac Formula)**.** *Let the n-dimensional process Z satisfy the SDE* $dZ_t = b(Z_t)dt + \sigma(Z_t)dW_t$ *where W is an m dimensional Brownian Motion. Given functions f, g, suppose* $v : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ *is* $C^2$*, bounded, and satisfies:*

$$\frac{\partial v}{\partial t} + \sum_i b^i\frac{\partial v}{\partial z^i} + \frac{1}{2}\sum_{i,j} a^{ij}\frac{\partial^2 v}{\partial z^i \partial z^j} = fV \tag{17}$$

*where* $a^{ij} = \sum_k \sigma^{ik}\sigma jk$ *with terminal conditions* $v(T, z) = g(z)$ *for all* $z \in \mathbb{R}^n$*. Let, for* $0 \leqslant t \leqslant T$*,*

$$M_t = v(t, Z_t)\exp\left\{-\int_0^t f(Z_s)ds\right\} \tag{18}$$

*Then, M is a local martingale. Further, if M is a true martingale, then*

$$v(t, z) = \mathbb{E}\left[\exp\left\{-\int_t^T f(Z_s)ds\right\}g(Z_T)\middle|Z_t = z\right] \tag{19}$$

*Proof.* By Itô's formula, and the fact that $v$ satisfies Equation (17), we have that

$$dM_t = \exp\left\{-\int_0^t f(Z_s)ds\right\}\sum_{i,j}\sigma^{ij}\frac{\partial v}{\partial z^i}dW^j \tag{20}$$

Hence, $M$ is a local martingale since it is a stochastic integral with respect to a Brownian Motion.
Now, if $M$ is a true martingale, the result follows from noting that $M_t = \mathbb{E}[M_T|\mathcal{F}_t]$ since

$$M_t = \mathbb{E}[M_T|\mathcal{F}_t] \tag{21}$$

$$v(t, Z_t)\exp\left\{-\int_0^t f(Z_s)ds\right\} = \mathbb{E}\left[g(Z_T)\exp\left\{-\int_0^T f(Z_s)ds\right\}\middle|\mathcal{F}_t\right] \tag{22}$$

$$\implies v(t, Z_t) = \exp\left\{\int_0^t f(Z_s)ds\right\}\mathbb{E}\left[g(Z_T)\exp\left\{-\int_0^T f(Z_s)ds\right\}\middle|\mathcal{F}_t\right] \tag{23}$$

$$\implies v(t, Z_t) = \mathbb{E}\left[g(Z_T)\exp\left\{-\int_t^T f(Z_s)ds\right\}\middle|\mathcal{F}_t\right] \tag{24}$$

Then, by the Tower Property of conditional expectation and noting that $\{Z_t = z\} \subset \mathcal{F}_t$:

$$\mathbb{E}[v(t, Z_t)|Z_t = z] = \mathbb{E}\left[\mathbb{E}\left[g(Z_T)\exp\left\{-\int_t^T f(Z_s)ds\right\}\middle|\mathcal{F}_t\right]\middle|Z_t = z\right] \tag{25}$$

$$v(t, z) = \mathbb{E}\left[g(Z_T)\exp\left\{-\int_t^T f(Z_s)ds\right\}\middle|Z_t = z\right] \tag{26}$$

$\square$

We can now apply this result to Equation (16) by realising that the function $f$ is constant and is equal to $r$, and our terminal function $g$ will be the payoff of the option we are pricing. In the context of pricing a call option, $g(S_T) = (S_T - K)^+$ where $K$ is the strike price of the option. Hence we get

$$V(t, S) = \mathbb{E}\left[\exp\left\{-\int_t^T rds\right\}g(S_T)\middle|S_t = S\right] \tag{27}$$

$$V(t, S) = \exp\{-r(T - t)\}\mathbb{E}\left[(S_T - K)^+|S_t = S\right] \tag{28}$$

$$V(t, S) = \exp\{-r(T - t)\}\int_{\mathbb{R}}\left(S\exp\left\{\left(r - \frac{\sigma^2}{2}\right)(T - t) + \sigma\sqrt{T - t}z\right\} - K\right)^+\phi(z)dz \tag{29}$$

where $\phi(z)$ is the pdf of a standard normal distribution: $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$. Solving this integral gives the Black-Scholes formula where

$$V(t,S) = S\Phi\left(-\frac{\log(K/S)}{\sigma\sqrt{T-t}} + \left(\frac{r}{\sigma} + \frac{\sigma}{2}\right)\sqrt{T-t}\right) - Ke^{-r(T-t)}\Phi\left(-\frac{\log(K/S)}{\sigma\sqrt{T-t}} + \left(\frac{r}{\sigma} - \frac{\sigma}{2}\right)\sqrt{T-t}\right) \tag{30}$$

where $\Phi$ is the cdf of the standard normal distribution. Taking $t = 0$ ie looking at the initial price of the option, we have

$$C(S,K,r,\sigma,T) = S\Phi(d_1) - Ke^{-rT}\Phi(d_2) \tag{31}$$

$$d_1 = -\frac{\log(K/S)}{\sigma\sqrt{T}} + \left(\frac{r}{\sigma} + \frac{\sigma}{2}\right)\sqrt{T} \tag{32}$$

$$d_2 = d1 - \sigma\sqrt{T} = -\frac{\log(K/S)}{\sigma\sqrt{T}} + \left(\frac{r}{\sigma} - \frac{\sigma}{2}\right)\sqrt{T} \tag{33}$$

By using the Put-Call parity formula, we can derive a similar closed-form solution for the price of a European Put:

$$P(S,K,r,\sigma,T) = Ke^{-rT}(1 - \Phi(d_2)) - S(1 - \Phi(d_1)) \tag{34}$$

where $d_1$ and $d_2$ are as above.

## 2.2 The Problem of Constant Volatility

In their seminal paper, Black & Scholes (1972) derived the above closed form solutions when working under 'ideal' conditions. The conditions they asserted were as follows

1. Frictionless markets, that is, there are no transaction costs
2. The spot interest rate is known and constant through time
3. Stocks pay no dividends
4. The option contract is European
5. The stock price follows a geometric Brownian Motion

Whilst work has been done to weaken some of these conditions, the main problem arises when we weaken the final assumption. If we make the concession that the dynamics of the underlying asset are not almost surely smooth, then the Black-Scholes framework breaks down. However, it is not obvious a priori that this is a worthwhile avenue to pursue. There is intellectual benefit from weakening assumptions in any model to better understand its limits, but what of the practical application. As a brief example, consider the historic volatilities of the largest component of the FTSE 100, Shell PLC (Ticker: SHEL). In Figure 1, we see the rolling one-year volatilities. Of striking note is the fact that contrary to the assumption of the Black-Scholes model, volatility does not remain as a constant throughout the time that data is available. Instead it varies in a 'random' way.

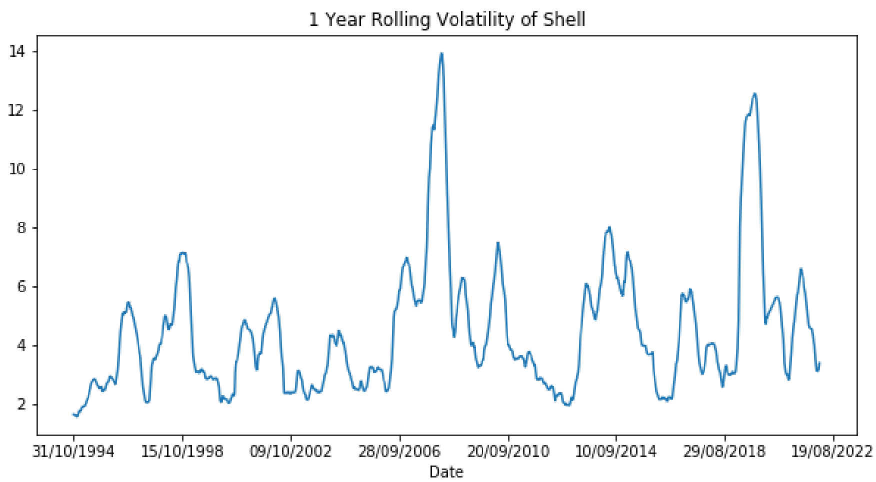


Figure 1: Above is the rolling 1 Year Volatility for Shell PLC. This is calculated as the standard deviation over one trading year (252 days) starting from the company's first listing in 30/10/1994. We note that over the past 20 years, volatility has not remained constant, which gives suitable practical use for pursuing stochastic volatility

Whilst this is just one example of a stock that exhibits this behaviour, this is a pattern that is seen in all stocks; there are periods where volatility is extremely high, and there are periods when the volatility is low. This motivates our discussion in the rest of this paper where we begin analysing models that allow volatility to be time varying in their own right

# 3 Heston Model

## 3.1 Model Set-Up

As noted in the previous section, the assumption that volatility remains constant in time does not accurately reflect the conditions seen in reality. This motivates a study of a class of models referred to as stochastic volatility models. As the name suggests, the underpinning of this family of models is to allow the volatility of the process to be evolve under its own process. There are a variety of different models that exist to model this including the Constant Elasticity of Variance (CEV) model and the GARCH model. However, in keeping with the historical development of these ideas, we will focus on the Heston Model. The model makes use of a slightly abridged version of Equation (1) as follows

$$dS_t = \mu S_t dt + \sqrt{\nu_t} S_t dW_t^1 \tag{35}$$

$$d\nu_t = \kappa(\theta - \nu_t)dt + \xi\sqrt{\nu_t} dW_t^2 \tag{36}$$

where $\mu$ is the percentage drift of the underlying asset, $\nu$ is understood as the instantaneous variance, $\theta$ is the long-run average of the variance of the price, $\kappa$ is the rate at which $\nu$ reverts to $\theta$, $\xi$ is the volatility of the volatility. The Brownian Motions $W_t^1$ and $W_t^2$ are correlated with a correlation coefficient of $\rho$. In order to initialise this process we require two initial values, namely $S_0$, the initial underlying price, and $\nu_0$, the initial variance. This model set-up comes from Heston (1993)
Upon first glance of this model, it can seem rather unwieldy, and there are naturally two questions that might arise. Firstly, what does it mean to have to correlated Brownian Motions? Secondly, we traditionally think of volatility as being positive in value, thus how can we ensure that $\nu_t > 0$ for all times $t$. We will now explore the answers to these questions

### 3.1.1 Correlated Brownian Motions

Suppose we are working in filtered probability space given by $(\Omega, \mathcal{F}, \mathbb{P})$. The goal will be to construct two correlated Brownian Motions with respect to the same filtration $\mathbb{F} = \bigcup_t \mathcal{F}_t$, where the filtration is assumed to satisfy the normal conditions. In order to do so we started with two uncorrelated Brownian Motions with respect to $\mathbb{F}$, we denote them by $W_t^1$ and $W_t^\perp$. Given that the two correlated Brownian Motions have correlation coefficient $\rho$, we can then create $W_t^2$ as follows

$$W_t^2 = \rho W_t^1 + \sqrt{1-\rho^2} W_t^\perp \tag{37}$$

We can then verify that $W_t^1$ and $W_t^2$ are indeed correlated by collecting and applying a few results.

Firstly, we note that $\mathbb{V}[W_t^1] = t$ and that

$$\mathbb{V}[W_t^2] = \rho^2 \mathbb{V}[W_t^1] + (1-\rho^2)\mathbb{V}[W_t^\perp] = \rho^2 t + (1-\rho^2)t = t \tag{38}$$

Hence, we find that $\sqrt{\mathbb{V}[W_t^1]} = \sqrt{\mathbb{V}[W_t^2]} = \sqrt{t}$.

Secondly, we can calculate the covariance between $W_t^1$ and $W_t^2$:

$$COV[W_t^1, W_t^2] = \rho COV[W_t^1, W_t^1] + \sqrt{1-\rho^2} COV[W_t^1, W_t^\perp] = \rho \mathbb{V}[W_t^1] = \rho t \tag{39}$$

Thus, we can explicitly calculate the correlation between $W_t^1$ and $W_t^2$ as follows

$$CORR[W_t^1, W_t^2] = \frac{COV[W_t^1, W_t^2]}{\sqrt{\mathbb{V}[W_t^1]}\sqrt{\mathbb{V}[W_t^2]}} = \frac{\rho t}{t} = \rho \tag{40}$$

as desired.

This construction also allows for us to efficiently create the two correlated Brownian Motions in Python since we create $W_t^1$ and $W_t^\perp$ and then for each time step we calculate $W_t^2 = \rho W_t^1 + \sqrt{1-\rho^2} W_t^\perp$

### 3.1.2 Non-Negativity of $\nu_t$

The next concern to address relates to the fact that the process

$$d\nu_t = \kappa(\theta - \nu_t)dt + \xi\sqrt{\nu_t} dW_t^2 \tag{41}$$

initialised with initial volatility $\nu_0$ is not a priori strongly positive. This process by which the volatility evolves is known as the Cox-Ingersoll-Ross (CIR) process and has its own rich history as an extension to the Vasicek model. As presented in Cox et al. (1985), this stochastic process was viewed as a method of understanding the evolution of the instantaneous interest rate. Intuitively, the dynamics of Equation (41) should prevent $\nu_t$ from becoming negative, since if $\nu_t$ is already close to 0, the standard deviation of the process, given by $\xi\sqrt{\nu_t}$ becomes small and so the evolution increasingly is driven by the drift term, which will be positive as long as $\theta > \nu_t$. However, formalising this hunch requires a bit of care. To prove this, we shall prove what is known as the Feller condition (cf. Feller (1951)) which states that as long as $2\kappa\theta > \xi^2$, then $\nu_t$ is bounded away from zero:

**Theorem 3.1** (Positivity of Volatility)**.** *Let $\nu_t$ evolve as in Equation (41). Then if the Feller condition, $2\kappa\theta > \xi^2$ is satisfied, then there exists a positive solution to Equation (41) on each time interval $t \in [0, \infty)$*

*Proof.* We will follow the proof outlined in Gikhman (2011) and we fill in some of the details. The proof can be broken down in to three steps.

1. Application of Itô's formula
2. Application of Grönwall's inequality
3. Application of Chebyshev's Inequality

Before we get started with these parts, let us first get set up. Let $\varepsilon > 0$ and denote $\tau_\varepsilon = \min\{t : \nu_t \leqslant \varepsilon\}$. It can be checked that $\tau_\varepsilon$ is a stopping time and measures the first time $\nu_t$ falls below $\varepsilon$. Further, let $\tau_\varepsilon \wedge t = \min\{\tau_\varepsilon, t\}$. To prove our result, we will show that as $\varepsilon \to 0$, $\mathbb{P}[\tau_\varepsilon \wedge t < t] \to 0$, ie , that the probability that $\nu_t$ is below $\varepsilon$ by time $t$ goes to zero. Now, define

$$m = \frac{2\kappa\theta - \xi^2}{\xi^2} \tag{42}$$

Step 1. We apply Itô's formula[1]in its integral form to the function $f(x) = x^{-m}$, which yields

$$\nu_{\tau_\varepsilon \wedge t}^{-m} = \nu_0^{-m} - \int_0^{\tau_\varepsilon \wedge t} m\xi\sqrt{\nu_s}\nu_s^{-(m+1)}dW_s - \int_0^{\tau_\varepsilon \wedge t} m(\kappa - \nu_s)\nu_s^{-(m+1)}ds + \frac{1}{2}\int_0^{\tau_\varepsilon \wedge t} m(m+1)\xi^2\nu_s\nu_s^{-(m+2)}ds \tag{43}$$

$$\nu_{\tau_\varepsilon \wedge t}^{-m} = \nu_0^{-m} + m\kappa\int_0^t \nu_{\tau_\varepsilon \wedge s}^{-m}ds - m\xi\int_0^t \nu_{\tau_\varepsilon \wedge s}^{-(m+\frac{1}{2})}dW_s + \left(\frac{m(m+1)}{2}\xi^2 - m\kappa\theta\right)\int_0^t \nu_{\tau_\varepsilon \wedge s}^{-(m+1)}ds \tag{44}$$

$$\nu_{\tau_\varepsilon \wedge t}^{-m} \leqslant \nu_0^{-m} + m\kappa\int_0^t \nu_{\tau_\varepsilon \wedge s}^{-m}ds \tag{45}$$

where we reach the last line by noting that the two-right most terms are of a smaller magnitude. Therefore, we can finish this step by noting

$$\mathbb{E}\left[\nu_{\tau_\varepsilon \wedge t}^{-m}\right] \leqslant \nu_0^{-m} + m\kappa\int_0^t \mathbb{E}[\nu_{\tau_\varepsilon \wedge s}^{-m}]ds \tag{46}$$

Step 2. Since $\nu_0^{-m}$ is constant we can apply Grönwall's inequality. A full statement and prove of Grönwall's inequality can be found in Appendix C. This tells us that

$$\mathbb{E}\left[\nu_{\tau_\varepsilon \wedge t}^{-m}\right] \leqslant \nu_0^{-m}e^{mkt} \tag{47}$$

Step 3. The result now follows from applying Chebyshev's inequality as follows

$$\begin{aligned}
\mathbb{P}[\tau_\varepsilon \leqslant t] &= \mathbb{P}[\nu_{\tau_\varepsilon}^m < \varepsilon^m] &(48)\\
&= \mathbb{P}[\nu_{\tau_\varepsilon}^{-m} \geqslant \varepsilon^{-m}] &(49)\\
&\leqslant \varepsilon^m\mathbb{V}[\nu_{\tau_\varepsilon}^{-m}] &(50)\\
&\leqslant \varepsilon^m\mathbb{E}[\nu_{\tau_\varepsilon}^{-m}] &(51)\\
&\leqslant \varepsilon^m\nu_0^{-m}e^{mkt} &(52)
\end{aligned}$$

which tends to 0 as $\varepsilon \to 0$

Hence if $2\kappa\theta > \xi^2$, that is, the Feller condition is satisfied, then the CIR process is positive for all $t$ as required ☐

[1]Itô's formula is usually quoted as follows: Let $X_t$ be an Itô process such that

$$dX_t = \alpha_t dW_t + \beta_t dt \tag{53}$$

Then if $f : \mathbb{R} \to \mathbb{R}$ is twice differentiable and for $(\alpha_t)_{t\geqslant 0}, (\beta_t)_{t\geqslant 0}$ previsible processes such that $\int_0^t \alpha_s^2 ds < \infty$ and $\int_0^t |\beta_s|ds < \infty$, then

$$df(X_t) = f'(X_t)dX_t + \frac{1}{2}f''(X_t)d\langle X\rangle_t \tag{54}$$

However, if we $X_t$ as an integral, ie

$$X_t = X_0 + \int_0^t \alpha_s dW_s + \int_0^t \beta_s ds \tag{55}$$

Then we can restate Itô's formula as saying

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\alpha_s dW_s + \int_0^t f'(X_s)\beta_s ds + \frac{1}{2}\int_0^t f''(X_s)\alpha_s^2 ds \tag{56}$$

### 3.2 Option Pricing

Much like we did previously, we shall now look at how we can use the Heston model to price European options. We shall follow a similar derivation to above by first deriving a PDE that we expect the value function to obey. The issue that we will face is that whilst we can derive a Feynman-Kac style PDE, we cannot explicitly use the Feynman-Kac formula due to the fact the derived PDE will have additional second derivative terms that we can not handle with this machinery. We will then look at ways we can then price options using Fourier pricing.

We shall consider a function $V : \mathbb{R}_{\geqslant 0} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $(t, S_t, \nu_t) \mapsto V(t, S_t, \nu_t)$. We then apply Itô's lemma to $V$:

$$\begin{aligned} dV(t,S_t,\nu_t) &= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS_t + \frac{\partial V}{\partial \nu}d\nu_t && (57)\\ &+ \frac{1}{2}\frac{\partial^2 V}{\partial t^2}\langle dt, dt\rangle + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\langle dS_t, dS_t\rangle + \frac{1}{2}\frac{\partial^2 V}{\partial \nu^2}\langle d\nu_t, d\nu_t\rangle && (58)\\ &+ \frac{\partial^2 V}{\partial t \partial S}\langle dt, dS_t\rangle + \frac{\partial^2 V}{\partial t \partial \nu}\langle dt, d\nu_t\rangle + \frac{\partial^2 V}{\partial \nu \partial S}\langle d\nu_t, dS_t\rangle && (59) \end{aligned}$$

We can greatly simplify the above by making use of our covariation identities by noting that $\langle dt, dt\rangle = 0$ and by noting that $\langle dt, dW_t^i\rangle = 0$ for $i = 1, 2$. This observation then implies that $\langle dt, d\nu_t\rangle = \langle dt, dS_t\rangle = 0$. Hence, the only terms that survive are the covariation of $S_t$ and $\nu_t$ as well as the variation of $S_t$ and $\nu_t$ themselves. We now derive them:

$$\begin{aligned} \langle d\nu_t, dS_t\rangle &= \langle \kappa(\theta - \nu_t)dt + \xi\sqrt{\nu_t}dW_t^2, \mu S_t dt + S_t\sqrt{\nu_t}dW_t^1\rangle && (60)\\ &= \xi S_t \nu_t \langle dW_t^1, dW_t^2\rangle && (61)\\ &= \rho\xi S_t \nu_t dt && (62) \end{aligned}$$

$$\begin{aligned} \langle d\nu_t, d\nu_t\rangle &= \langle \kappa(\theta - \nu_t)dt + \xi\sqrt{\nu_t}dW_t^2, \kappa(\theta - \nu_t)dt + \xi\sqrt{\nu_t}dW_t^2\rangle && (63)\\ &= \xi^2 \nu_t dt && (64) \end{aligned}$$

$$\begin{aligned} \langle dS_t, dS_t\rangle &= \langle \mu S_t dt + S_t\sqrt{\nu_t}dW_t^1, \mu S_t dt + S_t\sqrt{\nu_t}dW_t^1\rangle && (65)\\ &= S_t^2 \nu_t dt && (66) \end{aligned}$$

Putting these results into the Itô-Taylor expansion above, we see that

$$dV(t,S_t,\nu_t) = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS_t + \frac{\partial V}{\partial \nu}d\nu_t + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}S_t^2\nu_t dt + \frac{1}{2}\frac{\partial^2 V}{\partial \nu^2}\xi^2\nu_t dt + \frac{\partial^2 V}{\partial \nu \partial S}\rho\xi S_t\nu_t dt \tag{67}$$

Substituting in the dynamics for $dS_t$ and $d\nu_t$ we see that

$$\begin{aligned} dV(t,S_t,\nu_t) &= \left[\frac{\partial V}{\partial t} + \frac{1}{2}S_t^2\nu_t\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\xi^2\nu_t\frac{\partial^2 V}{\partial \nu^2} + \rho\xi S_t\nu_t\frac{\partial^2 V}{\partial \nu \partial S}\right]dt && (68)\\ &+ \frac{\partial V}{\partial S}(\mu S_t dt + \xi\sqrt{\nu_t}dW_t^1) + \frac{\partial V}{\partial \nu}(\kappa(\theta - \nu_t)dt + \xi\sqrt{\nu_t}dW_t^2) && (69)\\ dV(t,S_t,\nu_t) &= \left[\frac{\partial V}{\partial t} + \mu S_t\frac{\partial V}{\partial S} + \kappa(\theta - \nu_t)\frac{\partial V}{\partial \nu} + \frac{1}{2}S_t^2\nu_t\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\xi^2\nu_t\frac{\partial^2 V}{\partial \nu^2} + \rho\xi S_t\nu_t\frac{\partial^2 V}{\partial \nu \partial S}\right]dt && (70)\\ &+ S_t\sqrt{\nu_t}\frac{\partial V}{\partial S}dW_t^1 + \xi\sqrt{\nu_t}\frac{\partial V}{\partial \nu}dW_t^2 && (71) \end{aligned}$$

As above, we can now analyse the dynamics of $(e^{-rt}V(t,S_t,\nu_t))_t$. Again, we want this to be a local martingale and so we need the drift term to identically vanish. By the product rule we have:

$$\begin{aligned} d\left\{e^{-rt}V(t,S_t,\nu_t)\right\} &= e^{-rt}dV(t,S_t,\nu_t) - re^{-rt}V(t,S_t,\nu_t)dt && (72)\\ \implies d\left\{e^{-rt}V(t,S_t,\nu_t)\right\} &= \left[\frac{\partial V}{\partial t} + \mu S_t\frac{\partial V}{\partial S} + \kappa(\theta - \nu_t)\frac{\partial V}{\partial \nu} + \frac{1}{2}S_t^2\nu_t\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\xi^2\nu_t\frac{\partial^2 V}{\partial \nu^2} + \rho\xi S_t\nu_t\frac{\partial^2 V}{\partial \nu \partial S} - rV\right]e^{-rt}dt && (73)\\ &+ e^{-rt}S_t\sqrt{\nu_t}\frac{\partial V}{\partial S}dW_t^1 + e^{-rt}\xi\sqrt{\nu_t}\frac{\partial V}{\partial \nu}dW_t^2 && (74) \end{aligned}$$

Imposing the desire that we have a martingale gives us:

$$\frac{\partial V}{\partial t} + \mu S_t\frac{\partial V}{\partial S} + \kappa(\theta - \nu_t)\frac{\partial V}{\partial \nu} + \frac{1}{2}S_t^2\nu_t\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\xi^2\nu_t\frac{\partial^2 V}{\partial \nu^2} + \rho\xi S_t\nu_t\frac{\partial^2 V}{\partial \nu \partial S} = rV \tag{75}$$

At this stage, if we were to follow the analysis conducted with the derivation of the Black-Scholes option price, we would appeal to the Feynman-Kac formula outlined in Theorem 2.1. However, the structure of the dynamics in the Heston model means that we cannot we cannot compute the terminal value of $S_T$ directly and so applying Equation 19 directly would fail to work. Hope is not lost though, since there is an alternative method of computing the option price

### 3.2.1 A Foray into Fourier Pricing

The alternative for deriving a semi-analytic option price is to make a handful of substitutions and then apply a Fourier Transform. Whilst we will not got into the details of the Fourier Transform in the main body, we will use some fairly elementary properties that are fleshed out in greater detail in Appendix D. The outline of the method is follows, we will first make two changes of variables. Initially, we recast Equation 75 to consider the logarithm of the underlying. Next, instead of looking at the value at some time $t < T$, instead we consider the time to expiration only, which is given by $\tau = T - t$. Secondly, we use an ansatz to derive a PDE that two probability functions must satisfy. Thirdly, we use a Fourier Transform for the these probabilities and solve the resulting set of differential equations that arise. After piecing these solutions with our chosen ansatz, we can then have a way of estimating the price of a call in the Heston model. We can then repeat the analysis if desired to create a similar scheme to price a put as well.

Before we get started, we must first take a moment to highlight a slight departure from the way this method was initially proposed. In Heston (1993), equation 6 is the analogue of our equation 75. However, if one was to check carefully, they would note that the coefficient of $\frac{\partial U}{\partial \nu}$ in Heston contains an addition term of the form $\lambda(S, \nu, t)$. In Heston's exposition, he notes that this $\lambda$ represents the price of volatility risk. And whilst he goes on to make the assumption that this price is linear and continues with his analysis, we are not going to include this in our analysis. The reason is twofold, pragmatically, we do not want to include an additional modelling assumption in our model since we are strictly comparing the dynamics of the underling SDEs that govern option pricing between the Black-Scholes, Heston, and Merton models. The next logical question is whether dropping this volatility risk will change our analysis. This leads directly to our second reason for ignoring this term; appealing to Gatheral (2011), we make the assertion that the Heston process generates the risk-neutral measure so that the market price of volatility is set to zero. We could always recover the probability measure in which the volatility risk is present, but since our concern is the with the option prices that follows from this set-up we are free to choose this model in lieu of Heston's. We can now continue our analysis

*Step 1: Change of Variables*
We first recast Equation 75 in terms of the logarithm of the underlying asset price. Whilst the literature differs on whether to take the logarithm of the underlying or rather the logarithm of the ratio of the asset price and the strike price, we will proceed with the former in keeping with Heston's original paper. Setting $x = \log(S)$, we can use some principles from elementary calculus to show that

$$\frac{\partial V}{\partial S} = \frac{\partial x}{\partial S}\frac{\partial V}{\partial x} = \frac{1}{2}\frac{\partial V}{\partial x} \tag{76}$$

$$\frac{\partial^2 V}{\partial \nu \partial S} = \frac{\partial}{\partial \nu}\left\{\frac{\partial V}{\partial S}\right\} = \frac{\partial}{\partial \nu}\left\{\frac{1}{S}\frac{\partial V}{\partial x}\right\} = \frac{1}{S}\frac{\partial^2 V}{\partial \nu \partial x} \tag{77}$$

$$\frac{\partial^2 V}{\partial S^2} = \frac{\partial}{\partial S}\left\{\frac{\partial V}{\partial S}\right\} = \frac{\partial}{\partial S}\left\{\frac{1}{S}\frac{\partial V}{\partial x}\right\} = -\frac{1}{S^2}\frac{\partial V}{\partial x} + \frac{1}{S}\frac{\partial x}{\partial S}\frac{\partial^2 V}{\partial x^2} = -\frac{1}{S^2}\frac{\partial V}{\partial x} + \frac{1}{S^2}\frac{\partial^2 V}{\partial x^2} \tag{78}$$

Substituting these formulae into Equation 75 yields:

$$\frac{\partial V}{\partial t} + \mu\frac{\partial V}{\partial x} + \kappa(\theta - \nu_t)\frac{\partial V}{\partial \nu} + \frac{1}{2}\nu_t\left(-\frac{\partial V}{\partial x} + \frac{\partial^2 V}{\partial x^2}\right) + \frac{1}{2}\xi^2\nu_t\frac{\partial^2 V}{\partial \nu^2} + \rho\xi\nu_t\frac{\partial^2 V}{\partial \nu \partial x} = rV \tag{79}$$

$$\frac{\partial V}{\partial t} + \left(\mu - \frac{1}{2}\nu_t\right)\frac{\partial V}{\partial x} + \kappa(\theta - \nu_t)\frac{\partial V}{\partial \nu} + \frac{1}{2}\nu_t\frac{\partial^2 V}{\partial x^2} + \frac{1}{2}\xi^2\nu_t\frac{\partial^2 V}{\partial \nu^2} + \rho\xi\nu_t\frac{\partial^2 V}{\partial \nu \partial x} = rV \tag{80}$$

Next, we switch our time variable. In our set-up above, we were interested in pricing our option at some time $t$ that was before the maturity date of our contract $T$. We will change our perspective and instead time-index by our time to expiry. This will be denoted by $\tau = T - t$. This does not create any real structural change to Equation 75, other than giving us a minus sign for the time derivative as follows:

$$-\frac{\partial V}{\partial \tau} + \left(\mu - \frac{1}{2}\nu\right)\frac{\partial V}{\partial x} + \kappa(\theta - \nu)\frac{\partial V}{\partial \nu} + \frac{1}{2}\nu\frac{\partial^2 V}{\partial x^2} + \frac{1}{2}\xi^2\nu\frac{\partial^2 V}{\partial \nu^2} + \rho\xi\nu\frac{\partial^2 V}{\partial \nu \partial x} = rV \tag{81}$$

*Step 2: Ansatz*
In order to move forward with our analysis, we make the ansatz that the general structure of the price of the call option is similar to that of the Black-Scholes model. We note that since $x = \log(S)$, then $S = e^x$ so we use the following ansatz:

$$C(\tau, x, \nu) = e^x P_1(\tau, x, \nu) - Ke^{-r\tau}P_2(\tau, x, \nu) \tag{82}$$

We then have to compute the corresponding partial derivatives

$$\frac{\partial C}{\partial \tau} = e^x\frac{\partial P_1}{\partial \tau} + rKe^{-r\tau}P_2 - Ke^{-r\tau}\frac{\partial P_2}{\partial \tau} \qquad \frac{\partial C}{\partial x} = e^xP_1 + e^x\frac{\partial P_1}{\partial x} - Ke^{-r\tau}\frac{\partial P_2}{\partial x} \tag{83}$$

$$\frac{\partial^2 C}{\partial x^2} = e^xP_1 + 2e^x\frac{\partial P_1}{\partial x} + e^x\frac{\partial^2 P_1}{\partial x^2} - Ke^{-r\tau}\frac{\partial^2 P_2}{\partial x^2} \qquad \frac{\partial C}{\partial \nu} = e^x\frac{\partial P_1}{\partial \nu} - Ke^{-r\tau}\frac{\partial P_2}{\partial \nu} \tag{84}$$

$$\frac{\partial^2 C}{\partial \nu^2} = e^x\frac{\partial^2 P_1}{\partial \nu^2} - Ke^{-r\tau}\frac{\partial^2 P_2}{\partial \nu^2} \qquad \frac{\partial^2 C}{\partial \nu\partial x} = e^x\frac{\partial P_1}{\partial \nu} + e^x\frac{\partial^2 P_1}{\partial \nu\partial x} - Ke^{-r\tau}\frac{\partial^2 P_2}{\partial \nu\partial x} \tag{85}$$

Substituting the above five forms into Equation 81 and bracketing the derived equations in terms of $P_1$ and $P_2$, we see that:

$$\begin{aligned} r(e^xP_1(\tau,x,\nu) - Ke^{-r\tau}P_2(\tau,x,\nu)) = &-\left[e^x\frac{\partial P_1}{\partial \tau} + rKe^{-r\tau}P_2 - Ke^{-r\tau}\frac{\partial P_2}{\partial \tau}\right] + \left(\mu - \frac{1}{2}\nu\right)\left[e^xP_1 + e^x\frac{\partial P_1}{\partial x} - Ke^{-r\tau}\frac{\partial P_2}{\partial x}\right] \\ &+ \kappa(\theta-\nu)\left[e^x\frac{\partial P_1}{\partial \nu} - Ke^{-r\tau}\frac{\partial P_2}{\partial \nu}\right] + \frac{1}{2}\nu\left[e^xP_1 + 2e^x\frac{\partial P_1}{\partial x} + e^x\frac{\partial^2 P_1}{\partial x^2} - Ke^{-r\tau}\frac{\partial^2 P_2}{\partial x^2}\right] \\ &+ \frac{1}{2}\xi^2\nu\left[e^x\frac{\partial^2 P_1}{\partial \nu^2} - Ke^{-r\tau}\frac{\partial^2 P_2}{\partial \nu^2}\right] + \rho\xi\nu\left[e^x\frac{\partial P_1}{\partial \nu} + e^x\frac{\partial^2 P_1}{\partial \nu\partial x} - Ke^{-r\tau}\frac{\partial^2 P_2}{\partial \nu\partial x}\right] \end{aligned} \tag{86}$$

Cleaning this up we see that

$$\begin{aligned} &e^x\left[(\mu - r)P_1 - \frac{\partial P_1}{\partial \tau} + \left(\mu + \frac{1}{2}\nu\right)\frac{\partial P_1}{\partial x} + (\kappa(\theta-\nu) + \rho\xi\nu)\frac{\partial P_1}{\partial \nu} + \frac{1}{2}\nu\frac{\partial^2 P_1}{\partial x^2} + \rho\xi\nu\frac{\partial^2 P_1}{\partial \nu\partial x} + \frac{1}{2}\xi^2\nu\frac{\partial^2 P_1}{\partial \nu^2}\right] \\ &= \\ &Ke^{-r\tau}\left[-\frac{\partial P_2}{\partial \tau} + \left(\mu - \frac{1}{2}\nu\right)\frac{\partial P_2}{\partial x} + \kappa(\theta-\nu)\frac{\partial P_2}{\partial \nu} + \frac{1}{2}\nu\frac{\partial^2 P_2}{\partial x^2} + \rho\xi\nu\frac{\partial^2 P_2}{\partial \nu\partial x} + \frac{1}{2}\xi^2\nu\frac{\partial^2 P_2}{\partial \nu^2}\right] \end{aligned} \tag{87}$$

Setting $\mu = r$, much like we did in the Black-Scholes case, reduces the complexity of this equation by removing the $P_1$ term. Since this must hold for all $x$ and all $K$, we can disentangle these two equations by setting $K = 0$ to find an equation for $P_1$ and by setting $S \to 0$ to find the equation for $P_2$. These equations are typically summarised as below:

$$-\frac{\partial P_j}{\partial \tau} + (r + u_j\nu)\frac{\partial P_j}{\partial x} + (\kappa\theta - b_j\nu)\frac{\partial P_j}{\partial \nu} + \frac{1}{2}\nu\frac{\partial^2 P_j}{\partial x^2} + \rho\xi\nu\frac{\partial^2 P_j}{\partial \nu\partial x} + \frac{1}{2}\xi^2\nu\frac{\partial^2 P_j}{\partial \nu^2} = 0 \tag{88}$$

for $j = 1, 2$ where $u_1 = \frac{1}{2}$, $u_2 = -\frac{1}{2}$, and $b_1 = \kappa - \rho\xi$, $b_2 = \kappa$. The terminal conditions for these PDEs becomes

$$P_j(0, x, \nu; \log(K)) = \mathbb{1}\{x \geqslant \log(K)\} \tag{89}$$

*Step 3: Fourier Transform*
To solve Equation 91, we now introduce the Fourier Transform of $P_j$ which we define as follows:

$$\mathfrak{F}_x\{P_j(\tau,x,\nu)\}(\varphi) = \tilde{P}_j(\tau,\varphi,\nu) = \int_{-\infty}^{\infty} e^{-i\varphi x}P_j(\tau,x,\nu)dx \tag{90}$$

We note that, by our terminal condition given in Equation 92:

$$\tilde{P}_j(0,\varphi,\nu) = \int_{-\infty}^{\infty} e^{-i\varphi x}\mathbb{1}\{x \geqslant \log(K)\}dx = \int_{\log(K)}^{\infty} e^{-i\varphi x}dx = \frac{1}{i\varphi}e^{-i\varphi x}\Big|_{\infty}^{\log(K)} = \frac{Ke^{-i\varphi}}{i\varphi} \tag{91}$$

The aim now is to find an analogous PDE that the Fourier Transform of $P_j$ must satisfy. Whilst the intricacies of the following statements can be found in Appendix D, we use the following results about how derivatives and Fourier Transform interact. (we use $\mathfrak{F}_x$ to denote the Fourier Transform with respect to $x$)

$$\mathfrak{F}\left\{\frac{\partial P_j}{\partial \tau}\right\} = \frac{\partial \tilde{P}_j}{\partial \tau} \qquad \mathfrak{F}\left\{\frac{\partial P_j}{\partial \nu}\right\} = \frac{\partial \tilde{P}_j}{\partial \nu} \tag{92}$$

$$\mathfrak{F}\left\{\frac{\partial^2 P_j}{\partial \nu^2}\right\} = \frac{\partial^2 \tilde{P}_j}{\partial \nu^2} \qquad \mathfrak{F}\left\{\frac{\partial P_j}{\partial x}\right\} = i\varphi\tilde{P}_j \tag{93}$$

$$\mathfrak{F}\left\{\frac{\partial^2 P_j}{\partial x^2}\right\} = -\varphi^2\tilde{P}_j \qquad \mathfrak{F}\left\{\frac{\partial^2 P_j}{\partial x\partial \nu}\right\} = i\varphi\frac{\partial \tilde{P}_j}{\partial \nu} \tag{94}$$

We can now find a PDE for the Fourier Transform of $P_j$ by applying Equations (92)-(94) to Equation (88) as follows

$$-\frac{\partial \tilde{P}_j}{\partial \tau} + i\varphi(r + u_j\nu)\tilde{P}_j + (\kappa\theta - b_j\nu)\frac{\partial \tilde{P}_j}{\partial \nu} - \frac{1}{2}\varphi^2\nu\tilde{P}_j + i\varphi\rho\xi\nu\frac{\partial \tilde{P}_j}{\partial \nu} + \frac{1}{2}\xi^2\nu\frac{\partial^2 \tilde{P}_j}{\partial \nu^2} = 0 \tag{95}$$

$$\implies \left[i\varphi r\tilde{P}_j - \frac{\partial \tilde{P}_j}{\partial \tau} + \kappa\theta\frac{\partial \tilde{P}_j}{\partial \nu}\right] + \left[\left(i\varphi u_j - \frac{1}{2}\varphi^2\right)\tilde{P}_j + (i\varphi\rho\xi - b_j)\frac{\partial \tilde{P}_j}{\partial \nu} + \frac{1}{2}\xi^2\frac{\partial^2 \tilde{P}_j}{\partial \nu^2}\right]\nu = 0 \tag{96}$$

Before we power through solving Equation (96) to find functions for $\tilde{P}_j$, we will first take a moment to highlight some important characteristics of this equation that will help us in our solution. Firstly, since this equation must hold for all values of $\nu$, we can see that is must be the case that both square brackets must be identically zero. Secondly, we note that the PDE is linear in each of its coefficients, which would suggest some type of affine solution.

*Step IV: Second Ansatz*
Using our two observations above, we shall make use of the ansatz that $\tilde{P}_j$ will have an exponential affine form:

$$\tilde{P}_j(\tau,\varphi,\nu) = \tilde{P}_j(0,\varphi,\nu)\exp\{A(\tau,\varphi) + B(\tau,\varphi)\nu\} = \frac{K}{i\varphi}\exp\{A(\tau,\varphi) + B(\tau,\varphi)\nu - i\varphi\} \tag{97}$$

We can then calculate the partial derivatives of $\tilde{P}_j$ using this ansatz which gives:

$$\frac{\partial \tilde{P}_j}{\partial \tau} = \left(\frac{\partial A}{\partial \tau} + \nu\frac{\partial B}{\partial \tau}\right)\tilde{P}_j \tag{98}$$

$$\frac{\partial \tilde{P}_j}{\partial \nu} = B\tilde{P}_j \tag{99}$$

$$\frac{\partial^2 \tilde{P}_j}{\partial \nu^2} = B^2\tilde{P}_j \tag{100}$$

Thus, using these partial derivatives, we can alter Equation 102 to give:

$$\left[i\varphi r\tilde{P}_j - \left(\frac{\partial A}{\partial \tau} + \nu\frac{\partial B}{\partial \tau}\right)\tilde{P}_j + \kappa\theta B\tilde{P}_j\right] + \left[\left(i\varphi u_j - \frac{1}{2}\varphi^2\right)\tilde{P}_j + (i\varphi\rho\xi - b_j)B\tilde{P}_j + \frac{1}{2}\xi^2B^2\tilde{P}_j\right]\nu = 0 \tag{101}$$

$$\implies \left[i\varphi r - \frac{\partial A}{\partial \tau} + \kappa\theta B\right] + \left[\left(i\varphi u_j - \frac{1}{2}\varphi^2\right) + (i\varphi\rho\xi - b_j)B + \frac{1}{2}\xi^2B^2 - \frac{\partial B}{\partial \tau}\right]\nu = 0 \tag{102}$$

By noting as above that Equation (102) must hold for all values of $\nu$, we then have two first order differential equations to solve:

$$\frac{\partial A}{\partial \tau} = \kappa\theta B + i\varphi r \tag{103}$$

$$\frac{\partial B}{\partial \tau} = \left(i\varphi u_j - \frac{1}{2}\varphi^2\right) + (i\varphi\rho\xi - b_j)B + \frac{1}{2}\xi^2B^2 \tag{104}$$

with initial conditions $A(0) = B(0) = 0$

*Step V: Solving Differential Equations*
To solve the above system of differential equations is rather straight-forward, if a bit involved. The primary insight here is th note that Equation 110 is nothing more than a Riccati equation for $B$:

$$\frac{\partial B}{\partial \tau} = \frac{1}{2}\xi^2\left[B^2 + \frac{2(i\varphi\rho\xi - b_j)}{\xi^2}B + \frac{2(i\varphi u_j - \frac{1}{2}\varphi^2)}{\xi^2}\right] \tag{105}$$

$$\implies \frac{\frac{\partial B}{\partial \tau}}{\left[B^2 + \frac{2(i\varphi\rho\xi - b_j)}{\xi^2}B + \frac{2(i\varphi u_j - \frac{1}{2}\varphi^2)}{\xi^2}\right]} = \frac{1}{2}\xi^2 \tag{106}$$

By the quadratic formula, the denominator has roots

$$\underbrace{-\frac{i\varphi\rho\xi - b_j}{\xi^2}}_{:=\alpha} \pm \underbrace{\sqrt{\frac{(i\varphi\rho\xi - b_j)^2}{\xi^4} - \frac{2(i\varphi u_j - \frac{1}{2}\varphi^2)}{\xi^2}}}_{:=\beta} \tag{107}$$

Then Equation (106) becomes

$$\frac{\frac{\partial B}{\partial \tau}}{(B - (\alpha + \beta))(B - (\alpha - \beta))} = \frac{1}{2}\xi^2 \tag{108}$$

By performing a partial fraction decomposition we can show that

$$\frac{1}{2\beta}\left[\frac{1}{B-(\alpha+\beta)}-\frac{1}{B-(\alpha-\beta)}\right]\frac{\partial B}{\partial \tau}=\frac{1}{2}\xi^2 \tag{109}$$

$$\left[\frac{1}{B-(\alpha+\beta)}-\frac{1}{B-(\alpha-\beta)}\right]\frac{\partial B}{\partial \tau}=\xi^2\beta \tag{110}$$

$$\implies \int_\tau^0\left[\frac{1}{B-(\alpha+\beta)}-\frac{1}{B-(\alpha-\beta)}\right]\frac{\partial B}{\partial s}ds=\int_\tau^0 \xi^2\beta ds \tag{111}$$

$$\implies \log\left(\frac{B(s,\varphi)-(\alpha+\beta)}{B(s,\varphi)-(\alpha-\beta)}\right)\Bigg|_\tau^0=-\xi^2\beta\tau \tag{112}$$

$$\implies \log\left(\frac{\alpha+\beta}{\alpha-\beta}\right)-\log\left(\frac{B(\tau,\varphi)-(\alpha+\beta)}{B(\tau,\varphi)-(\alpha-\beta)}\right)=-\xi^2\beta\tau \tag{113}$$

$$\implies \frac{B(\tau,\varphi)-(\alpha+\beta)}{B(\tau,\varphi)-(\alpha-\beta)}=\frac{\alpha+\beta}{\alpha-\beta}e^{\xi^2\beta\tau} \tag{114}$$

$$\implies \left[1-\frac{\alpha+\beta}{\alpha-\beta}e^{\xi^2\beta\tau}\right]B(\tau,\varphi)=(\alpha+\beta)-(\alpha+\beta)e^{\xi^2\beta\tau} \tag{115}$$

$$\implies B(\tau,\varphi)=\frac{(\alpha+\beta)\left[1-e^{\xi^2\beta\tau}\right]}{1-\frac{\alpha+\beta}{\alpha-\beta}e^{\xi^2\beta\tau}} \tag{116}$$

Whilst we could attempt to work on from here, typically, the function $B$ is not given in this form. We shall state the final form here, but work through the rest of this part in Appendix E. Working through the details yields the following result

$$B(\tau,\varphi)=\frac{(e^{M\tau}-1)(i\varphi\rho\xi+b_j-M)}{\xi^2(1-Ne^{M\tau})} \tag{117}$$

$$M=\sqrt{(i\varphi\rho\xi-b_j)^2-\xi^2(2iu_j\varphi-\varphi^2)} \tag{118}$$

$$N=\frac{i\varphi\rho\xi+b_j-M}{i\varphi\rho\xi+b_j+M} \tag{119}$$

With this, we can now find a form for $A$:

$$\int_\tau^0\frac{\partial A}{\partial s}ds=\int_\tau^0 i\varphi r ds+\kappa\theta\int_\tau^0 B(s,\varphi)ds \tag{120}$$

$$A(\tau,\varphi)=i\varphi r\tau+\kappa\theta\int_\tau^0 B(s,\varphi)ds \tag{121}$$

Again, at this point, we yield our space to Appendix E, where we get the final result that

$$A(\tau,\varphi)=i\varphi r\tau+\frac{\kappa\theta}{\xi^2}\left[(i\varphi\rho\xi+b_j-M)\tau-2\log\left(\frac{1-N}{1-Ne^{M\tau}}\right)\right] \tag{122}$$

### 3.2.2 Putting The Pieces Together

Let us pause briefly to summarise where we are. We have now found a closed form solution for the Fourier Transform of the pseudo probabilities as follows

$$\tilde{P}_j(\tau,\varphi,\nu)=\frac{K}{i\varphi}\exp\{A(\tau,\varphi)+B(\tau,\varphi)\nu-i\varphi\} \tag{123}$$

$$A(\tau,\varphi)=i\varphi r\tau+\frac{\kappa\theta}{\xi^2}\left[(i\varphi\rho\xi+b_j-M)\tau-2\log\left(\frac{1-N}{1-Ne^{M\tau}}\right)\right] \tag{124}$$

$$B(\tau,\varphi)=\frac{(e^{M\tau}-1)(i\varphi\rho\xi+b_j-M)}{\xi^2(1-Ne^{M\tau})} \tag{125}$$

$$M=\sqrt{(i\varphi\rho\xi-b_j)^2-\xi^2(2iu_j\varphi-\varphi^2)} \tag{126}$$

$$N=\frac{i\varphi\rho\xi+b_j-M}{i\varphi\rho\xi+b_j+M} \tag{127}$$

$$u_1=\frac{1}{2} \tag{128}$$

$$u_2=-\frac{1}{2} \tag{129}$$

$$b_1=\kappa-\rho\xi \tag{130}$$

$$b_2=\kappa \tag{131}$$

To recover the original probabilities, we now must invert our Fourier Transform (details in Appendix D), where in particular we use the Gil-Pelaez (1951) inversion to yield that

$$P_j(\tau, x, \nu) = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \mathfrak{Re}\left\{\frac{\exp\{A(\tau,\varphi) + B(\tau,\varphi)\nu + i\varphi(x - \log(K))\}}{i\varphi}\right\} d\varphi \tag{132}$$

Note that whilst the Gil-Pelaez inversion applies to characteristic functions, since we are applying the inversion to the Fourier Transform of a probability function, we are indeed working with characteristic functions. We have not explicitly called the $P_j$ since a priori we do not know whether do satisfy the requirements of a probability density function or not

### 3.2.3 Some Final Notes

Before we move on, there a few things that are worth noting at this stage.

Firstly, whilst these may seem daunting calculations to complete by hand, they can be done efficiently on a computer. Recall that the price of the European Call option is then $C = SP_1 - Ke^{-rT}P_2$ where we then evaluate the $P_j$ at $\tau = T$ since $\tau$ captures the time to expiration

Secondly, we have skipped over some details with regards to the complex logarithm that arise in the calculation of $A(\tau, \varphi)$. Primarily, we have made the implicit assumption that for all $0 \leqslant \tau \leqslant T$, that the complex logarithm in $A$ is continuous. If we were dealing with the real logarithm, this would amount to verifying that the argument of the logarithm is strictly positive. However, with the complex logarithm, the task of verifying continuity becomes more subtle due to the fact that we must take care with the branch we choose in defining the complex logarithm. Broadly speaking, there are three ways of handling this issue.

1. Assume that this is not a problem in our case and move on
2. A priori specify the principal branch on which the complex logarithm is defined (Schöbel & Zhu 1999, Lee et al. 2004)
3. Keep track of the winding number of the complex logarithm and add the appropriate multiple of $2\pi$ when it is time to evaluate (Sepp 2003)

Whilst all three have their places in literature, in this paper we shall keep things simple and take the first method

Thirdly, as mentioned earlier, we can then also find the price of a put option in this framework. This is done simply by leveraging put-call parity : let $\text{CALL}_0$ be the initial price of the call, let $\text{PUT}_0$ be the initial price of the put. Further, let both options start with an underlying sport price of $S_0$ and strike price $K$, and let both options have expiration $T$, then

$$\text{CALL}_0 - \text{PUT}_0 = S_0 - Ke^{-rT} \tag{133}$$

Hence, we see that

$$\begin{aligned} \text{PUT}_0 &= \text{CALL}_0 - S_0 + Ke^{-rT} && (134)\\ \implies \text{PUT}_0 &= Ke^{-rT}(1 - P_2(T, x, \nu)) - S_0(1 - P_1(T, x, \nu)) && (135) \end{aligned}$$

## 3.3 Implementation

### 3.3.1 Numerical Approximation of Process

In order to create sample paths in the Heston model we, need a way to generate the paths of both $S_t$ and $\nu_t$. In order to do this, we first specify a time mesh over which we will be generating sample paths. If we are creating paths over time interval $[0, T]$, we specify a mesh that is dictated by $N \in \mathbb{Z}_{>0}$ such that $\Delta t = \frac{T}{N}$. We can then use an explicit formula to generate $S_{t+\Delta t}$ from $(S_t, \nu_t)$ and we use an Euler-Maruyama scheme to generate $\nu_{t+\Delta t}$ from $\nu_t$ as follows

$$\begin{aligned} S_{t+\Delta t} &= S_t e^{\left(r - \frac{1}{2}\nu_t\right)\Delta t + \sqrt{\nu_t}(W^1_{t+\Delta t} - W^1_t)} && (136)\\ \nu_{t+\Delta t} &= \nu_t + \kappa(\theta - \nu_t)\Delta t + \xi\sqrt{\nu_t}(W^2_{t+\Delta t} - W^2_t) && (137) \end{aligned}$$

By leveraging the definition of the Brownian Motion, and letting $z, z^{\perp} \sim \mathcal{N}(0,1)$ such that $COV[z, z^{\perp}] = 0$ we can bring the above system into a more computer friendly iteration scheme

$$\begin{aligned} S_{t+\Delta t} &= S_t e^{\left(r - \frac{1}{2}\nu_t\right)\frac{T}{N} + \sqrt{\nu_t}\sqrt{\frac{T}{N}}z} && (138)\\ \nu_{t+\Delta t} &= \nu_t + \kappa(\theta - \nu_t)\frac{T}{N} + \xi\sqrt{\nu_t}\sqrt{\frac{T}{N}}(\rho z + \sqrt{1-\rho^2}z^{\perp}) && (139) \end{aligned}$$

Below in Figure 2, we generate some sample paths using the Heston framework. Whilst this is more of a demonstration to help familiarise ourselves with the dynamics, we can note a few key observations. Firstly, the choice of process for volatility was done in order to preserving the mean reverting nature seen in reality. As shown below, the volatility processes eventually approach the 'long-term volatility average', $\theta$, despite starting at some other $\nu_0$. This mean-reverting behaviour is one of the main reasons the CIR process is chosen for the volatility. Secondly, we note that visually, there is some correlation between the randomness in $S_t$ and the randomness in $\nu_t$ (we can see that there are some periods of time when both processes move in the same or opposite directions). Recalling our structural equations, we can see that this is governed by our choice of $\rho$, or rather, the correlation between the Brownian Motions that drive the randomness in $S_t$ and $\nu_t$. In our specific case, we chose a value for $\rho$ to be close to 1 so we would expect to see more periods where the movements in $S_t$ and $\nu_t$ are the in the same direction


Figure 2: Here are some sample paths of the asset price (left panel) and the volatility (right panel). These three samples have been initialised with the following parameters: $T = 10, N = 100000, S_0 = 100, \nu_0 = 0.1, r = 0.35, \kappa = 1.5, \theta = 0.7, \xi = 0.2, \rho = 0.7$. In the left panel, the purple dotted line represents $S_0$, and in these cases we see that in each our simulations, the asset price rose. In the right panel, the black dotted line represents the value of $\nu_0$ and the red dotted line represents $\theta$. We can see that whilst all the volatilities start at $\nu_0$, by the end of the time window, the are all around the value of $\theta$ demonstrating the mean reversion in volatility that is seen empirically

#### 3.3.2 Numerically Pricing Options

In order to price the options, we need to compute this Fourier Transform integral. Whilst this can be tricky in Python, we use the open-source code that was created by Nicola Cantarutti[2]. The idea is to appeal to the characteristic function of the Heston model and apply a numerical approximation for the integral required

### 3.4 Comparison to Black-Scholes

We can now compare the option prices that are derived under the Heston model and compare them to their Black-Scholes counterparts. Intuitively, we expect that the Heston price should be higher than the Black-Scholes price since the stochastic volatility component of the Heston model creates enough risk that the option write ought to charge a higher price. We can see this in Figure 3

[2] `https://github.com/cantaro86`

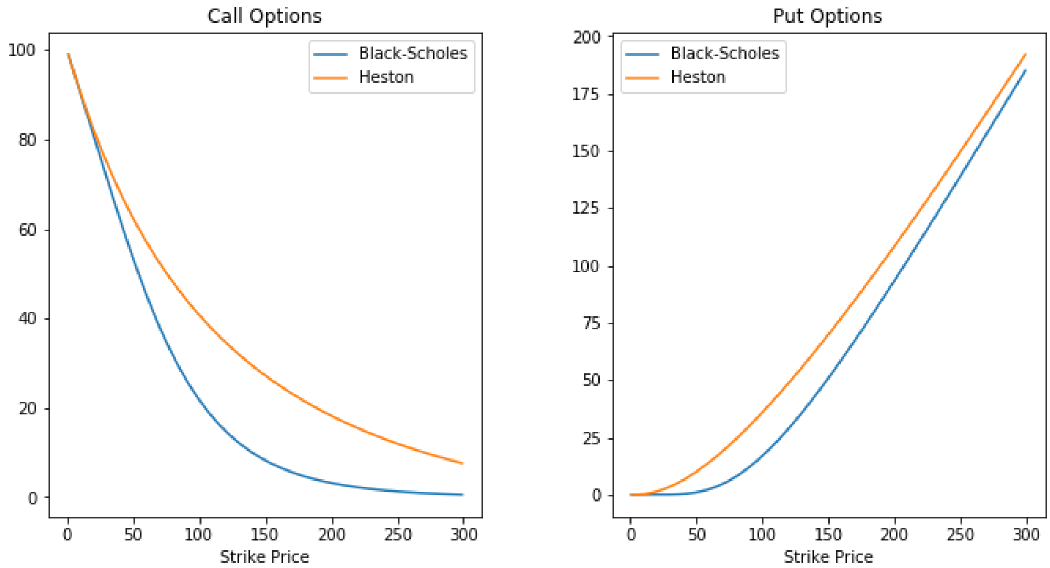


Figure 3: To begin with, we choose the parameters $T = 1, \nu_0 = 0.1, r = 0.05, \kappa = 3, \theta = 0.5, \xi = 0.8, \rho = 0.7, S_0 = 100$. In order to remain consistent between both model, we use the long term mean of the variance, $\theta$ as the variance that is used in the Black-Scholes model. We can see with this model configuration, the Heston option price remains above the Black-Scholes option price, indicating that there is enough movement in the volatility to require the option writer to give a higher price

# 4 Merton Model

We now turn our attention to a different extension of the Black-Scholes model that comes from Merton (1976). The model proposed in this paper works to extend the Black-Scholes set-up by weakening the assumption that the underlying asset evolves in a continuous manner, ie, in a way that is not fully captured by the Geometric Brownian Motion. Intuitively, this could result in cases when new (understood in the mathematical sense as not being previsible) information about the underlying is presented. Since this information could not be predicted ahead of time, it could not be priced into the options that are derived from the underlying. In reality, some examples of this could be the announcement of earning reports or other pieces of information that might not be strictly be financial, such as news stories etc. (In his paper, Merton refers to these types of events as abnormal vibrations in price). The key idea is to model these abnormal vibrations in a way that respects the martingale structure that is required in the analysis, which leads to inclusion of the a 'Poisson-driven' process.

Before we introduce the rest of this section, a few notes are worth making. Firstly, in the main body of the text, we will take for granted the existence and construction of a Poisson process. Details of this can be found in Appendix F. Secondly, in his paper, Merton allows for the jump-process to be of a general form before identifying the case in which the jumps are lognormally distributed. We shall begin with this as a modelling assumption, but mention in passing the cases where alternate distributions for the jumps might lead to other insights

The rest of this section begins with an introduction to the model set-up before deriving the formula for an option price in this model. Finally we explain the implementation

## 4.1 Model Set-Up

The defining equation for the underlying process is given by

$$dS_t = (\alpha - \lambda k)S_t dt + \sigma S_t dW_t + (y_t - 1)S_t dN_t \tag{140}$$

where $\alpha$ is the mean return of the asset, $\sigma$ is the volatility of the asset conditional on no jumps occurring, $W_t$ is a standard Brownian Motion process, $N_t$ is a Poisson process with intensity $\lambda$. $y_t$ is a nonnegative random variable that controls the size of the price jumps, and we will assume that it is lognormally distributed, ie $Y_t := \log(y_t) \sim \mathcal{N}(\mu, \delta^2)$[3]where $\mu, \delta$ are the mean and variance of the log of the jumps. Further, we define $k = \mathbb{E}[y_t - 1] = e^{\mu + \frac{1}{2}\delta^2} - 1$

The structure of the stochastic differential equation varies by practitioner. One of the big areas where the literature diverges comes to the structure of the drift term. We choose this form in order to retain the fact the the expectation of $dS_t$ is indeed $\alpha S_t dt$ since

$$\mathbb{E}\left[\frac{dS_t}{S_t}\right] = \mathbb{E}[(\alpha - \lambda k)dt] + \mathbb{E}[\sigma dW_t] + \mathbb{E}[(y_t - 1)dN_t] = (\alpha - \lambda k)dt + 0 + \lambda \underbrace{\left(e^{\mu + \frac{1}{2}\delta^2} - 1\right)}_{=k} dt = \alpha dt \tag{141}$$

However, this is not the only way to specify this process; we can insist that the drift term is $\alpha dt$, but to compensate, the Brownian Motion must then have a drift term, that is $W_t \sim \mathcal{N}\left(-\frac{\lambda k t}{\sigma}, t\right)$

### 4.1.1 Solving the Model

In the case of the Merton model, we can explicitly derive a closed-form solution for the asset price which we will do here

**Proposition 4.1** (Closed Form Solution of the Merton Jump Model). *Let $S_t$ obey the stochastic differential equation outlined in Equation (140). Then*

$$S_t = S_0 \exp\left\{\left(\alpha - \frac{1}{2}\sigma^2 - \lambda k\right)t + \sigma W_t + \sum_{i=1}^{N_t} \log(y_t)\right\} \tag{142}$$

*Proof.* The proof is essentially is just an application of Itô's lemma. We will apply this to $f(t, S_t) = \log(S_t)$. We use a version of Itô's formula that incorporates the Poisson-process as well (Tankov 2003) which states that

$$df(t, S_t) = \frac{\partial f}{\partial t}dt + b_t \frac{\partial f}{\partial S}dt + \frac{1}{2}\sigma^2 \frac{\partial^f}{\partial S^2}dt + \sigma \frac{\partial f}{\partial S}dW_t + [f(t, S_t + \Delta S_t) - f(t, S_t)] \tag{143}$$

[3]We note that if $\log(y_t) \sim \mathcal{N}(\mu, \delta^2)$, then $y_t \sim \mathcal{N}\left(e^{\mu + \frac{1}{2}\delta^2}, e^{2\mu + \delta^2}(e^{\delta^2} - 1)\right)$. This can be derived directly from the definition of the lognormal distribution

where $b_t$ corresponds to the drift term, $\sigma$ corresponds to the volatility term, and $\Delta S$ comes from the Poisson jump. By straight forward application we get that

$$\begin{aligned} d\log(S_t) &= \frac{\partial \log(S_t)}{\partial t}dt + (\alpha - \lambda k)S_t\frac{\partial \log(S_t)}{\partial S_t}dt + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 \log(S_t)}{\partial S_t^2}dt \\ &\quad + \sigma S_t \frac{\partial \log(S_t)}{\partial S_t}dW_t + [\log(y_t S_t) - \log(S_t)] \end{aligned} \tag{144}$$

$$d\log(S_t) = (\alpha - \lambda k)S_t \frac{1}{S_t}dt + \frac{1}{2}\sigma^2 S_t^2\left(-\frac{1}{S_t^2}\right)dt + \sigma S_t \frac{1}{S_t}dW_t + \log(y_t) \tag{145}$$

$$d\log(S_t) = \left(\alpha - \lambda k - \frac{1}{2}\sigma^2\right)dt + \sigma dW_t + \log(y_t) \tag{146}$$

$$\implies \log(S_t) - \log(S_0) = \left(\alpha - \lambda k - \frac{1}{2}\sigma^2\right)t + \sigma(W_t - W_0) + \sum_{i=1}^{N_t}\log(y_i) \tag{147}$$

$$\implies S_t = S_0 \exp\left\{\left(\alpha - \frac{1}{2}\sigma^2 - \lambda k\right)t + \sigma W_t + \sum_{i=1}^{N_t}\log(y_t)\right\} \tag{148}$$

□

## 4.2 Option Pricing

We will now use this closed form solution to derive the price of a European option in the Merton model. We will be following the derivations outlined in Merton (1976) and Matsuda (2004)

**Proposition 4.2.** *For an asset obeying the dynamics as outlined in Equation (140), the option price $V^M$ is given by*

$$V_t^M(S, K, r, \sigma, T, \lambda, \mu, \delta) = \sum_{n=0}^{\infty}\frac{e^{-\lambda}(\lambda)^n}{n!}V_t^{B\text{-}S}(S, K, \tilde{r}, \tilde{\sigma}, T) \tag{149}$$

*where $\tilde{r} = r - \lambda k + \frac{n(\mu + \frac{1}{2}\delta^2)}{2}$ and $\tilde{\sigma} = \sqrt{\sigma^2 + \frac{n\delta^2}{T-t}}$, and $V_t^{B-S}$ is the time $t$ Black-Scholes price*

*Proof.* The idea of the proof is to condition on the number of jumps that will occur between time $t$ and expiration at time $T$. For ease of expression, we will drop the arguments when we refer to $V^{\mathrm{M}}$ and $V^{\text{B-S}}$. We know that

$$V_t^{\mathrm{M}} = e^{-r(T-t)}\mathbb{E}\left[g(S_T)|\mathcal{F}_t\right] \tag{150}$$

where $g(\cdot)$ is the payoff function that the option is based on. Given that we have a closed form solution for $S_T$, we see that

$$V_t^{\mathrm{M}} = e^{-r(T-t)}\mathbb{E}\left[g\left(S_t \exp\left\{\left(r - \frac{1}{2}\sigma^2 - \lambda k\right)(T-t) + \sigma(W_T - W_t) + \sum_{i=N_t}^{N_T} Y_i\right\}\right)\right] \tag{151}$$

At this point, we condition on the number of jumps that occur between $t$ and $T$ as follows:

$$V_t^{\mathrm{M}} = e^{-r(T-t)}\sum_{n=0}^{\infty}\mathbb{P}[N_T - N_t = n]\mathbb{E}\left[g\left(S_t \exp\left\{\left(r - \frac{1}{2}\sigma^2 - \lambda k\right)(T-t) + \sigma\sqrt{T-t}z + \sum_{i=1}^{n} Y_i\right\}\right)\right] \tag{152}$$

$$V_t^{\mathrm{M}} = e^{-r(T-t)}\sum_{n=0}^{\infty}\frac{e^{-\lambda(T-t)}(\lambda(T-t))^n}{n!}\mathbb{E}\left[g\left(S_t \exp\left\{\left(r - \frac{1}{2}\sigma^2 - \lambda k\right)(T-t) + \sigma\sqrt{T-t}z + \sum_{i=1}^{n} Y_i\right\}\right)\right] \tag{153}$$

Recalling that $Y_i = \log(y_i) \sim \mathcal{N}(\mu, \delta^2)$, we note that $\sum_{i=1}^{n} Y_i \sim \mathcal{N}(n\mu, n\delta^2)$. Thus, the exponent is normally distributed and so we can rearragne it in such a way that keeps the mean and variance the same ie

$$V_t^{\mathrm{M}} = e^{-r(T-t)}\sum_{n=0}^{\infty}\frac{e^{-\lambda(T-t)}(\lambda(T-t))^n}{n!}\mathbb{E}\left[g\left(S_t \exp\left\{\left(r - \frac{1}{2}\sigma^2 - \lambda k\right)(T-t) + n\mu + \sqrt{\sigma^2(T-t) + n\delta^2}z\right\}\right)\right] \tag{154}$$

Letting $\tilde{\sigma} = \sqrt{\sigma^2 + \frac{n\delta^2}{T-t}}$, and noting that this implies that $\sigma^2 = \tilde{\sigma}^2 - \frac{n\delta^2}{T-t}$ ,we simplify to

$$V_t^{\mathrm{M}} = e^{-r(T-t)}\sum_{n=0}^{\infty}\frac{e^{-\lambda(T-t)}(\lambda(T-t))^n}{n!}\mathbb{E}\left[g\left(S_t \exp\left\{\left(r - \frac{1}{2}\left(\tilde{\sigma}^2 - \frac{n\delta^2}{T-t}\right) - \lambda k\right)(T-t) + n\mu + \tilde{\sigma}\sqrt{T-t}z\right\}\right)\right] \tag{155}$$

$$V_t^{\mathrm{M}} = e^{-r(T-t)}\sum_{n=0}^{\infty}\frac{e^{-\lambda(T-t)}(\lambda(T-t))^n}{n!}\mathbb{E}\left[g\left(S_t \exp\left\{\left(r - \lambda k + \frac{n(\mu + \frac{1}{2}\delta^2)}{T-t} - \frac{1}{2}\tilde{\sigma}^2\right)(T-t) + \tilde{\sigma}\sqrt{T-t}z\right\}\right)\right] \tag{156}$$

Recalling that

$$V^{\text{B-S}} = e^{-r(T-t)}\mathbb{E}\left[g\left(S_t \exp\left\{(r - \frac{1}{2}\sigma^2)(T-t) + \sigma\sqrt{T-t}z\right\}\right)\right] \tag{157}$$

We see that by setting $\tilde{r} = r - \lambda k + \frac{n(\mu+\frac{1}{2}\delta^2)}{T-t}$ and taking care with the change of measure that occurs from moving between the Merton and Black-Scholes pricing

$$V_t^{\text{M}} = \sum_{n=0}^{\infty} \frac{e^{-\tilde{\lambda}(T-t)}(\tilde{\lambda}(T-t))^n}{n!} V_t^{\text{B-S}}(S, K, \tilde{r}, \tilde{\sigma}, T) \tag{158}$$

where $\tilde{\lambda} = \lambda(1+k) = \lambda e^{\mu+\frac{1}{2}\delta^2}$ (this slight amendment comes from the fact that under the Black-Scholes model, the risk-neutral measure is such that the discounted geometric Brownian Motion is a martingale, whereas in the Merton model, requiring that $(e^{-rt}S_t)_{t\geqslant 0}$ is a martingale gives an additional $\lambda k$ term which we absorb into the the Poisson part of the equation) $\square$

## 4.3 Implementation

### 4.3.1 Numerical Approximation of Process

In order to simulate the underlying process that accounts for the jumps that characterise the Merton model, we can note that if we split our time window $[0, T]$ into $n$ smaller windows[4]such that $\Delta t = \frac{T}{n}$ we can use the approximation

$$S_{t+\Delta t} = S_t \exp\{\left(r - \frac{1}{2}\sigma^2\right)\Delta t + \sigma(W_{t_\Delta t} - W_t) + \sum_{i=N_t}^{N_{t+\Delta t}} Y_i\} \tag{159}$$

The next question that follows is how we go about creating the Poisson process in the first place. We choose to use a Bernoulli approximation scheme. For some small timestep $\Delta t$, the probability that we have a jump in our process is a Bernoulli trial with probability of success being $\lambda dt$. Notice that since the probability that our process jumps by two is $o(h)$ so we have only two possible outcomes. Further, since each increment is independent, if we view our time period as a succession of smaller time windows, in which we have independent Bernoulli trials, we then can approximate the overall Poisson process by a binomial approximation. Thus, it remains to show that this Bernoulli approximation converges in distribution to the desired Poisson process. We will now show this

**Theorem 4.3.** *For some $T > 0$ and $k \in \mathbb{R}_{>0}$, create a time windows $\tau_n(i) = \left(\frac{iT}{n}, \frac{(i+1)T}{n}\right)$. Then*

$$\lim_{n\to\infty} \mathbb{P}[N(t) = k] = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \tag{160}$$

*Proof.* Since we have a sequence of independent Bernoulli trials, we are working with a Binomial distribution, ie $N(t) \sim \text{Bin}\left(n, \frac{\lambda t}{n}\right)$ Hence,

$$\lim_{n\to\infty} \mathbb{P}\left[N(t) = k\right] = \lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda t}{n}\right)^k \left(1 - \frac{\lambda t}{n}\right)^{n-k} \tag{161}$$

$$= (\lambda t)^k \lim_{n\to\infty} \binom{n}{k} n^{-k} \left(1 - \frac{\lambda t}{n}\right)^n \left(1 - \frac{\lambda t}{n}\right)^{-k} \tag{162}$$

We will now consider each term in the limit one by one:

$$\lim_{n\to\infty} \left(1 - \frac{\lambda t}{n}\right)^n = e^{-\lambda t} \tag{163}$$

$$\lim_{n\to\infty} \left(1 - \frac{\lambda t}{n}\right)^{-k} = 1 \tag{164}$$

$$\lim_{n\to\infty} \binom{n}{k} n^{-k} = \frac{1}{k!} \lim_{n\to\infty} \frac{n!}{(n-k)!n^k} \tag{165}$$

$$= \frac{1}{k!} \lim_{n\to\infty} \frac{n(n-1)\cdots(n-k+2)(n-k+1)}{n^k} \tag{166}$$

$$= \frac{1}{k!} \lim_{n\to\infty} \frac{o(n^k)}{n^k} \tag{167}$$

$$= \frac{1}{k!} \tag{168}$$

[4]In previous sections discussing the implementation of these stochastic differential equations, we used $N$ to denote the number of windows we split out time $T$ into. However, since we are using $N_t$ in this model to denote the Poisson process, we instead use $n$ to denote the number of time windows we are using

Hence, putting everything together, we get that

$$\lim_{n\to\infty} \mathbb{P}\left[N(t)=k\right] = \frac{(-\lambda t)^k e^{\lambda t}}{k!} \tag{169}$$

□

This gives us a way of creating the Poisson process and the rest of the underlying process can then be created in a straightforward manner. In Figure 4 we can see this in action; we can create our sample paths and in doing so we can see how the jump discontinuities influence the evolution of the asset price

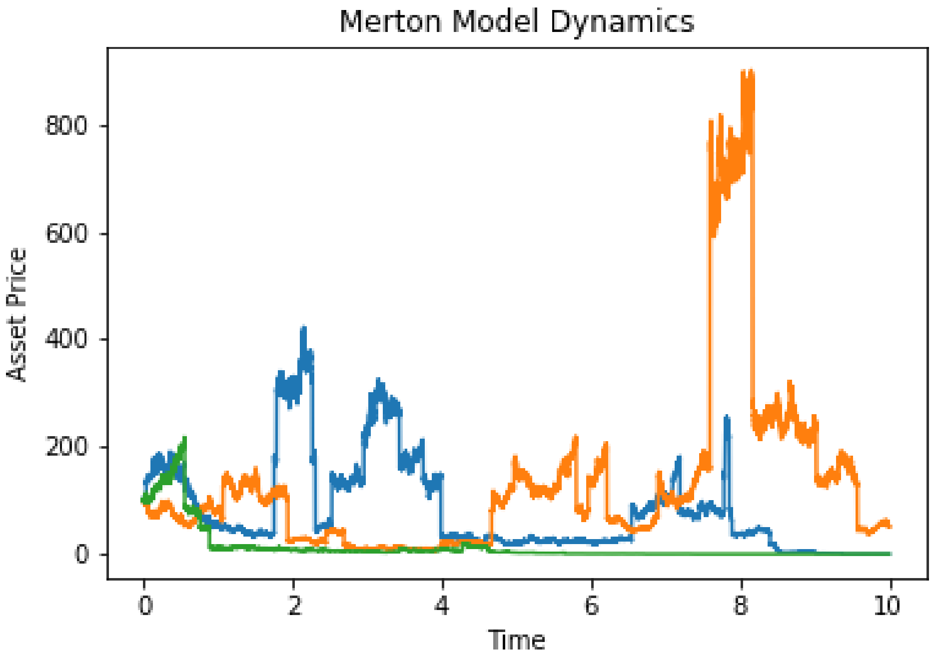


Figure 4: Here we show three sample paths of the underlying asset using the Merton model. We use the following parameters: $T = 10, n = 100000, S_0 = 100, \alpha = 1.2, \sigma = 0.5, \lambda - 3, \mu = 0, \delta = 0.8$. The most striking element of this of these sample path is the fact that we can see these jump discontinuities that characterise the Merton model

#### 4.3.2 Numerically Pricing Options

There is not too much that is required to be added in this section, except for one small detail that is used in the implementation of the option price. Recall, that the theoretical option price is given by

$$V_t^{\mathrm{M}} = \sum_{n=0}^{\infty} \frac{e^{-\tilde{\lambda}(T-t)}(\tilde{\lambda}(T-t))^n}{n!} V_t^{\text{B-S}}(S, K, \tilde{r}, \tilde{\sigma}, T) \tag{170}$$

where $\tilde{\lambda} = \lambda(1+k) = \lambda e^{\mu + \frac{1}{2}\delta^2}$. In practice, this summation to $\infty$ does not seem practical. Instead there are two ways of handling this:

1. We can arbitrarily set the upper limit for the number of jumps we choose to condition upon. Since computers are very efficient at this type of calculation, we could set the upper limit for the number of jumps to be relatively high and use this as a 'good' proxy for $\infty$
2. We can compute the magnitude of the contribution to the sum for each value $k$, and once the contribution is less than some defined tolerance parameter, we can terminate the process

Whilst either method works fine, it seems simpler for our analysis to set the upper limit of jumps since we will be varying parameters such as strike price and volatility, and this can have unintended impacts upon the magnitude of each contribution. In our code, we do set the upper limit to be 40 jumps, though we could go higher if we wish

### 4.4 Comparison to Black-Scholes

In this section, we briefly look at the comparison between the option price derived under the Black-Scholes and Merton models. Intuitively, we would expect the option price under the Merton model to be higher than the Black-Scholes option price. This is because allowing for jumps in the underlying increases the volatility of the underlying asset price. This is a result we can show analytically as follows

**Proposition 4.4.** *For some time $t \in [0, T]$, the volatility of $S_t$ in the Merton model is strictly greater than the volatility of $S_t$ in the Black-Scholes model.*

*Proof.* It will be convenient for us to work with $\log(S_t)$, however since the logarithm is a monotonic transform of $S_t$, proving this statement for $\log(S_t)$ is sufficient. Recall that

$$\log(S_t^{\text{B-S}}) = \log(S_0) + \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma W_t \tag{171}$$

$$\log(S_t^{\text{M}}) = \log(S_0) + \left(\alpha - -\frac{1}{2}\sigma^2 - \lambda k\right) t + \sigma W_t + \sum_{i=1}^{N_t} Y_i \tag{172}$$

We can see then, by using the definition of a Brownian Motion that

$$\mathbb{V}\left[\log(S_t^{\text{B-S}})\right] = \sigma^2 \mathbb{V}\left[W_t\right] = \sigma^2 t \tag{173}$$

Doing the same for the Merton asset price yields

$$\begin{aligned}
\mathbb{V}\left[\log(S_t^{\text{M}})\right] &= \sigma^2 \mathbb{V}\left[W_t\right] + \mathbb{V}\left[\sum_{i=1}^{N_t} Y_i\right] && (174)\\
&= \sigma^2 t + \mathbb{E}\left[\mathbb{V}\left[\sum_{i=1}^{N_t} Y_i \middle| N_t = n\right]\right] + \mathbb{V}\left[\mathbb{E}\left[\sum_{i=1}^{N_t} Y_i \middle| N_t = n\right]\right] && (175)\\
&= \sigma^2 t + \mathbb{E}\left[N_t \delta^2\right] + \mathbb{V}\left[N_t \mu\right] && (176)\\
&= \sigma^2 t + \lambda \delta^2 t + \lambda \mu^2 t && (177)\\
&> \sigma^2 t && (178)\\
&= \mathbb{V}\left[\log(S_t^{\text{B-S}})\right] && (179)
\end{aligned}$$

Where we use the Law of Total Variance[5] to calculate the variance of the Poisson process, and we conclude by noting that $\lambda > 0$ and $\delta^2, \mu^2 \geqslant 0$ □

Since the underlying asset in the Merton model has a higher volatility, we would expect the option price to be higher in this setting since the option writer is exposing themselves to a higher level of risk than in the Black-Scholes universe. Indeed if we run simulations, we see that this line of reasoning seems to hold true; Figure 5 shows that the Merton option will always cost more than the Black-Scholes option

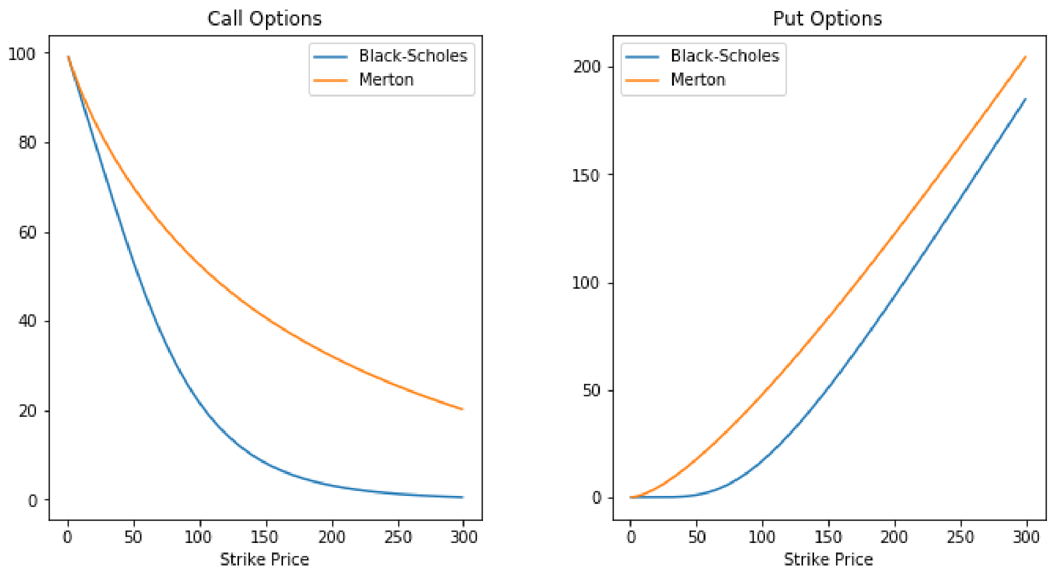


Figure 5: Here we see the price for call and put options in both the Black-Scholes and Merton universes. For this specific example, we initialised all options prices with $S_0 = 100, T = 1, r = 0.05, \sigma = 0.5$. Then, specifically for the Merton models, we set $\lambda = 2, \mu = 0.3, \delta = 0.2$. As mentioned above the crucial takeaway is that for all strike prices, the Merton option has a higher value than the Black-Scholes option

[5] The Law of Total Variance states that if $X$ and $Y$ are random variables on the same probability space, and $Y \in L^2$, then

$$\mathbb{V}[Y] = \mathbb{E}\left[\mathbb{V}\left[Y|X\right]\right] + \mathbb{V}\left[\mathbb{E}\left[Y|X\right]\right] \tag{180}$$

# A Derivation of Geometric Brownian Motion Solution

In this appendix we aim to find a closed form solution for the stochastic differential equation of the form

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{181}$$

We first note that $d\langle S\rangle_t = \sigma^2 S_t^2 dt$. Hence we can use Itô's lemma with $f(S) = \log(S)$ to show that

$$d(\log(S_t)) = \frac{1}{S_t} dS_t - \frac{1}{2}\frac{1}{S_t^2} d\langle S\rangle_t \tag{182}$$

$$d(\log(S_t)) = \frac{1}{S_t}\left(\mu S_t dt + \sigma S_t dW_t\right) - \frac{1}{2}\frac{1}{S_t^2}\sigma^2 S_t^2 dt \tag{183}$$

$$d(\log(S_t)) = \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dW_t \tag{184}$$

Then, integrating both sides yields

$$\int_0^t d(\log(S)) = \int_0^t \mu - \frac{1}{2}\sigma^2 dt + \int_0^t \sigma dW_t \tag{185}$$

$$\log(S_t) - \log(S_0) = \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma(W_t - W_0) \tag{186}$$

$$\frac{S_t}{S_0} = \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma W_t\right\} \tag{187}$$

$$S_t = S_0 \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma W_t\right\} \tag{188}$$

# B Alternate Derivation of the Black-Scholes Formula

Recall that our problem is to find a way of pricing an option where we know the payoff of the option at its maturity. In the main body, we appealed to the Feynman-Kac theorem, since it gives a way from moving from the Black-Scholes PDE to the Black-Scholes formula. However, this is not the only way of deriving the Black-Scholes PDE. Whilst the method provided in the main body gives allows a full solution to be derived, it does not provide an intuitive understanding for what is going on. The derivation in this appendix will hopefully serve to make the derivation seem a little more intuitive. The idea here is to create a riskless, self-financing portfolio and examine the resulting equation when we impose these requirements.

We let $\Pi$ be our self-financing portfolio that is comprised of one option and some amount $\Delta$ of the underlying asset. Requiring that the portfolio is riskless, whilst is has a technical definition, in our case amounts to saying that the value of the portfolio remains stable with regard to the changes in the price of the underlying asset. Thus, we have that

$$\Pi_t = V(t, S_t) + \Delta S_t \tag{189}$$

$$\implies d\Pi_t = dV(t, S_t) + \Delta dS_t \tag{190}$$

We can then apply Itô's lemma to $dV$ and use the fact that $dS_t$ evolves under the geometric Brownian Motion to see that

$$d\Pi_t = \left(\frac{\partial V}{\partial t} + \mu S_t \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} + \Delta \mu S_t\right) dt + \left(\sigma S_t \frac{\partial V}{\partial S} + \Delta \sigma S_t\right) dW_t \tag{191}$$

Requiring that the portfolio is riskless amounts to requiring that the 'randomness' in the portfolio is zero, that is, that the $dW_t$ term vanishes. In effect this implies that $\Delta = -\frac{\partial V}{\partial S}$ (this is known as the delta of the portfolio, hence the choice of symbol). Imposing this, Equation 169 becomes

$$d\Pi_t = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2}\right) dt \tag{192}$$

We also require that our portfolio attains the risk free rate, which means that $d\Pi_t = r\Pi dt = r\left(V(t, S_t) - \frac{\partial V}{\partial S} S_t\right) dt$. Therefore, combining these two expressions for $d\Pi_t$ yields

$$\left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2}\right) dt = r\left(V(t, S_t) - \frac{\partial V}{\partial S} S_t\right) dt \tag{193}$$

$$\implies \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} = r\left(V(t, S_t) - \frac{\partial V}{\partial S} S_t\right) \tag{194}$$

$$\implies \frac{\partial V}{\partial t} + r S_t \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} = rV \tag{195}$$

as required. Other derivations are available (Rouah 2020)

# C Grönwall's Inequality

In this appendix, we state and prove Grönwall's inequality. We will state and prove the integral version of the statement as that is often the one that is used the most frequently

**Theorem C.1** (Grönwall's Inequality)**.** *Suppose there are constants $a \in \mathbb{R}$ and $b > 0$ such that the bounded Borel function $f$ satisfies, for all $t \geqslant 0$*

$$f(t) \leqslant a + b \int_0^t f(s)ds \tag{196}$$

*Then*

$$f(t) \leqslant ae^{bt} \tag{197}$$

*Proof.* By the assumption that $f$ is a bounded Borel function, we can apply Fubini's theorem to conclude that

$$\begin{aligned} \int_{s=0}^t \int_{u=0}^s be^{b(t-s)} f(u)duds &= \int_{u=0}^t \int_{s=0}^t be^{b(t-s)} f(u)dsdu && (198) \\ &= \int_{u=0}^t (e^{b(t-u)} - 1) f(u)du && (199) \end{aligned}$$

Hence we have that

$$\begin{aligned} \int_0^t f(s)ds &= \int_0^t e^{t-s} \left( f(s) - b \int_0^s f(u)du \right) ds && (200) \\ &\leqslant \int_0^t e^{b(t-s)} ads && (201) \\ &= \frac{a}{b}(e^{bt} - 1) && (202) \end{aligned}$$

Thus, putting this back into Equation 174, we arrive at the result:

$$\begin{aligned} f(t) &\leqslant a + b \int_0^t f(s)ds && (203) \\ &\leqslant a + a(e^{bt} - 1) && (204) \\ &\leqslant ae^{bt} && (205) \end{aligned}$$

□

# D Fourier Transform and Analysis

In this appendix, we cover some key results in Fourier Analysis. We start with some definitions, before proving some results about how the Fourier Transform interacts with derivatives. We finish this appendix by stating and proving the Gil-Pelaez inversion Theorem

## D.1 Definitions

In its most general definition, we can define, for a function $f \in L^1(\mathbb{R}^n)$, the Fourier Transform $\tilde{f} : \mathbb{R}^n \to \mathbb{C}$ by

$$\mathfrak{F}\{f\}(\varphi) = \tilde{f}(\varphi) = \int_{\mathbb{R}^n} f(x)e^{-i\langle x,\varphi\rangle}dx \tag{206}$$

where $x, \varphi \in \mathbb{R}^n$

We note that the Fourier Transform is well defined since $\left|f(x)e^{-i\langle x,\varphi\rangle}\right| \leqslant |f(x)|$ and $f \in L^1(\mathbb{R}^n)$

## D.2 Fourier Transform and Derivatives

Suppose now that $f$ takes in two inputs. We will be consider how the derivatives of $f$ interact with the Fourier Transform. First, we define $f$ by

$$f : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \tag{207}$$

$$(x, t) \mapsto f(x, t) \tag{208}$$

We will now consider the Fourier Transform of $\frac{\partial f}{\partial x}$. The trick here is to use integration by parts

$$\mathfrak{F}\left\{\frac{\partial f}{\partial x}\right\}(\varphi) = \int_{\mathbb{R}} \frac{\partial f}{\partial x}e^{-ix\varphi}dx \tag{209}$$

$$= f(x)e^{-ix\varphi}\Big|_{-\infty}^{\infty} + i\varphi\int_{\mathbb{R}} f(x)e^{-ix\varphi}dx \tag{210}$$

$$= i\varphi\mathfrak{F}\{f\}(\varphi) \tag{211}$$

Hence we can also derive that $\mathfrak{F}\left\{\frac{\partial^2 f}{\partial x^2}\right\}(\varphi) = -\varphi^2\mathfrak{F}\{f\}(\varphi)$

Next, we assess the Fourier Transform of $\frac{\partial f}{\partial t}$. In this case, the idea is to commute the derivative and the integral as follows

$$\mathfrak{F}\left\{\frac{\partial f}{\partial t}\right\}(\varphi) = \int_{\mathbb{R}} \frac{\partial f}{\partial t}e^{-ix\varphi}dx \tag{212}$$

$$= \int_{\mathbb{R}} \lim_{h\to 0}\frac{f(x, t+h) - f(x, t)}{h}e^{-ix\varphi}dx \tag{213}$$

$$= \lim_{h\to 0}\frac{1}{h}\int_{\mathbb{R}}(f(x, t+h) - f(x, t))e^{-ix\varphi}dx \tag{214}$$

$$= \lim_{h\to 0}\frac{\mathfrak{F}\{f(x, t+h)\}(\varphi) - \mathfrak{F}\{f(x, t)\}(\varphi)}{h} \tag{215}$$

$$= \frac{\partial}{\partial t}\{\mathfrak{F}\{f\}(\varphi)\} = \frac{\partial\tilde{f}}{\partial t} \tag{216}$$

Hence we see that if we have a derivative term that is 'unrelated' to the Fourier Transform, then Fourier Transform of the derivative is just the Fourier Transform of the derivative

## D.3 Gil-Pelaez Inversion

We now state and prove the Gil-Pelaez Inversion theorem. The idea here is to leverage the connection between the probability density function of a random variable and its characteristic function. In the probability setting, the characteristic function is just the Fourier Transform of the probability density function. Let $f_X(x)$ be the probability density function for some random variable $x$, and let the corresponding cumulative density function be $F_X(x)$. Let the characteristic function for $X$ be given by

$$\phi_X(u) = \mathbb{E}\left[e^{iuX}\right] = \int_{-\infty}^{\infty} e^{iux}f_X(x)dx \tag{217}$$

**Theorem D.1** (Gil-Pelaez (1951)). *The cumulative density function for some random variable $X$ is given by*

$$F_X(x) = \frac{1}{2} + \frac{1}{2\pi}\int_0^\infty \frac{e^{iux}\phi_X(-u) - e^{-iux}\phi_X(u)}{iu}du \tag{218}$$

*Proof.* First, consider the step-function

$$\mathrm{sign}(y-x) = \frac{2}{\pi}\int_0^\infty \frac{\sin(u(y-x))}{u}du = \begin{cases} -1 & y < x \\ 0 & y = x \\ 1 & y > x \end{cases} \tag{219}$$

We note that

$$\begin{aligned} \int_{\mathbb{R}} \mathrm{sign}(y-x) f_X(y)dy &= -\int_{-\infty}^x f_X(y)dy + \int_x^\infty f_X(y)dy && (220)\\ &= -\left[F_X(x) - F_X(-\infty)\right] + \left[F_X(\infty) - F_X(x)\right] && (221)\\ &= 1 - 2F_X(x) && (222) \end{aligned}$$

Now, for some constants $a, b \in \mathbb{R}$ such that $a < b$

$$\begin{aligned} \frac{1}{\pi}\int_a^b \frac{e^{iux}\phi_X(-u) - e^{-iux}\phi_X(u)}{iu}du &= \frac{1}{\pi}\int_a^b \int_{\mathbb{R}} \frac{e^{-iu(y-x)} - e^{iu(y-x)}}{iu} f_X(y)dydu && (223)\\ &= -\frac{2}{\pi}\int_a^b \int_{\mathbb{R}} \frac{\sin(u(y-x))}{u} f_X(y)dydu && (224)\\ &= \left[-\int_{\mathbb{R}} f_X(y)dy\right]\left[\frac{2}{\pi}\int_a^b \frac{\sin(u(y-x))}{u}du\right] && (225) \end{aligned}$$

Where we exchange limits by appealing to Fubini's Theorem. Then letting $a \to 0$ and $b \to \infty$ we have that

$$\begin{aligned} \frac{1}{\pi}\int_0^\infty \frac{e^{iux}\phi_X(-u) - e^{-iux}\phi_X(u)}{iu}du &= \left[-\int_{\mathbb{R}} f_X(y)dy\right]\left[\lim_{\substack{a\to 0\\ b\to\infty}} \frac{2}{\pi}\int_a^b \frac{\sin(u(y-x))}{u}du\right] && (226)\\ &= -\int_{\mathbb{R}} \mathrm{sign}(y-x) f_X(y)dy && (227)\\ &= 2F_X(x) - 1 && (228) \end{aligned}$$

Hence we arrive at the conclusion that

$$F_X(x) = \frac{1}{2} + \frac{1}{2\pi}\int_0^\infty \frac{e^{iux}\phi_X(-u) - e^{-iux}\phi_X(u)}{iu}du \tag{229}$$

□

**Corollary D.2.** *We can alternatively express Equation 202 as follows*

$$F_X(x) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathfrak{Re}\left\{\frac{e^{-iux}\phi_X(u)}{iu}\right\}du \tag{230}$$

*Proof.* This follows by simple application of the fact that

$$\frac{\phi_X(u) + \phi_X(-u)}{2} = \mathfrak{Re}\{\phi_X(u)\} \tag{231}$$

And noting that $\phi_X(u) = \overline{\phi_X(y)}$ where the overline denotes complex conjugation

$$\begin{aligned} F_X(x) &= \frac{1}{2} + \frac{1}{2\pi}\int_0^\infty \frac{e^{iux}\phi_X(-u) - e^{-iux}\phi_X(u)}{iu}du && (232)\\ &= \frac{1}{2} + \frac{1}{2\pi}\int_0^\infty \frac{\overline{-e^{-iux}\phi_X(u)} - e^{-iux}\phi_X(u)}{iu}du && (233)\\ &= \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathfrak{Re}\left\{\frac{e^{-iux}\phi_X(u)}{iu}\right\}du && (234) \end{aligned}$$

□

In our application, we are interested in the 'complementary' cumulative density function, ie we want to find probabilities of the form $F_X(x)^c = \mathbb{P}[X \geqslant x]$. This can be done by noting

$$F_X^c(x) = 1 - F_X(x) \tag{235}$$

$$= 1 - \left( \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \mathfrak{Re} \left\{ \frac{e^{-iux}\phi_X(u)}{iu} \right\} du \right) \tag{236}$$

$$= \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathfrak{Re} \left\{ \frac{e^{-iux}\phi_X(u)}{iu} \right\} du \tag{237}$$

# E Derivation of $A(\tau, \varphi)$ and $B(\tau, \varphi)$

Recall that so far we have that

$$B(\tau, \varphi) = \frac{(\alpha + \beta)\left[1 - e^{\xi^2 \beta \tau}\right]}{1 - \frac{\alpha+\beta}{\alpha-\beta} e^{\xi^2 \beta \tau}} \tag{238}$$

We shall simplify this considerably. To do so, we also recall that

$$\alpha = -\frac{i\varphi\rho\xi - b_j}{\xi^2} \tag{239}$$

$$\beta = \sqrt{\frac{(i\varphi\rho\xi - b_j)^2}{\xi^4} - \frac{2(i\varphi u_j - \frac{1}{2}\varphi^2)}{\xi^2}} \tag{240}$$

We first begin by rewriting $\beta$ as follows

$$\beta = \frac{1}{\xi^2} \underbrace{\sqrt{(i\varphi\rho\xi - b_j)^2 - \xi^2(2iu_j\varphi - \varphi^2)}}_{:=M} \tag{241}$$

By cleaning up the expression for $B$ we see that

$$B(\tau, \varphi) = \frac{(\alpha^2 - \beta^2)\left[1 - e^{\xi^2 \frac{M}{\xi^2} \tau}\right]}{(\alpha - \beta) - (\alpha + \beta) e^{\xi^2 \frac{M}{\xi^2} \tau}} \tag{242}$$

$$= \frac{(\alpha^2 - \beta^2)\left[1 - e^{M\tau}\right]}{(\alpha - \beta) - (\alpha + \beta) e^{M\tau}} \tag{243}$$

We then note that

$$\alpha^2 - \beta^2 = \left(-\frac{i\varphi\rho\xi - b_j}{\xi^2}\right)^2 - \left(\frac{(i\varphi\rho\xi - b_j)^2}{\xi^4} - \frac{2(i\varphi u_j - \frac{1}{2}\varphi^2)}{\xi^2}\right) \tag{244}$$

$$= \frac{2(i\varphi u_j - \frac{1}{2}\varphi^2}{\xi^2} \tag{245}$$

Substituting this back into $B$ yields

$$B(\tau, \varphi) = \frac{1 - e^{M\tau}}{\xi^2} \frac{2i\varphi u_j - \varphi^2}{(\alpha - \beta) - (\alpha + \beta) e^{M\tau}} \tag{246}$$

We now simplify the denominator:

$$\xi^2\left[(\alpha - \beta) - (\alpha + \beta) e^{M\tau}\right] = \xi^2\left[\left(-\frac{i\varphi\rho\xi - b_j}{\xi^2} - \frac{M}{\xi^2}\right) - \left(-\frac{i\varphi\rho\xi - b_j}{\xi^2} + \frac{M}{\xi^2}\right) e^{M\tau}\right] \tag{247}$$

$$= (-i\varphi\rho\xi - b_j - M) - (-i\varphi\rho\xi - b_j + M) e^{M\tau} \tag{248}$$

so

$$B(\tau, \varphi) = \frac{(1 - e^{M\tau})(2i\varphi u_j - \varphi^2)}{(-i\varphi\rho\xi - b_j - M) - (-i\varphi\rho\xi - b_j + M) e^{M\tau}} \tag{249}$$

$$= \frac{(1 - e^{M\tau}) \frac{2i\varphi u_j - \varphi^2}{-i\varphi\rho\xi - b_j - M}}{1 - \frac{-i\varphi\rho\xi - b_j + M}{-i\varphi\rho\xi - b_j - M} e^{M\tau}} \tag{250}$$

$$= \frac{(1 - e^{M\tau}) \frac{\varphi^2 - 2i\varphi u_j}{i\varphi\rho\xi + b_j + M}}{1 - \frac{i\varphi\rho\xi + b_j - M}{i\varphi\rho\xi + b_j + M} e^{M\tau}} \tag{251}$$

Letting

$$N = \frac{i\varphi\rho\xi + b_j - M}{i\varphi\rho\xi + b_j + M} \tag{252}$$

Gives that

$$B(\tau, \varphi) = \frac{(1 - e^{M\tau}) \frac{\varphi^2 - 2i\varphi u_j}{i\varphi\rho\xi + b_j + M}}{1 - N e^{M\tau}} \tag{253}$$

We then turn our attention to the numerator:

$$\frac{\varphi^2 - 2i\varphi u_j}{i\varphi\rho\xi + b_j + M} = \frac{\left[\varphi^2 - 2i\varphi u_j\right][i\varphi\rho\xi + b_j - M]}{[i\varphi\rho\xi + b_j + M][i\varphi\rho\xi + b_j - M]} \tag{254}$$

$$= \frac{\left[\varphi^2 - 2i\varphi u_j\right][i\varphi\rho\xi + b_j - M]}{(i\varphi\rho\xi + b_j)^2 - M^2} \tag{255}$$

$$= \frac{\left[\varphi^2 - 2i\varphi u_j\right][i\varphi\rho\xi + b_j - M]}{\xi^2(2iu_j\varphi - \varphi^2)} \tag{256}$$

$$= -\frac{i\varphi\rho\xi + b_j - M}{\xi^2} \tag{257}$$

Thus we get a final form for $B$ as follows

$$B(\tau, \varphi) = \frac{(e^{M\tau} - 1)(i\varphi\rho\xi + b_j - M)}{\xi^2(1 - Ne^{M\tau})} \tag{258}$$

$$M = \sqrt{(i\varphi\rho\xi - b_j)^2 - \xi^2(2iu_j\varphi - \varphi^2)} \tag{259}$$

$$N = \frac{i\varphi\rho\xi + b_j - M}{i\varphi\rho\xi + b_j + M} \tag{260}$$

With $B$ calculated, we turn our attention to solving for $A$

$$A(\tau, \varphi) = i\varphi r\tau + \kappa\theta \int_\tau^0 B(s, \varphi) ds \tag{261}$$

$$= i\varphi r\tau + \kappa\theta \int_\tau^0 \frac{(e^{Ms} - 1)(i\varphi\rho\xi + b_j - M)}{\xi^2(1 - Ne^{Ms})} ds \tag{262}$$

$$= i\varphi r\tau + \frac{\kappa\theta}{\xi^2}(i\varphi\rho\xi + b_j - M) \int_\tau^0 \frac{e^{Ms} - 1}{1 - Ne^{Ms}} ds \tag{263}$$

In particular, we will focus on dealing with the integral term as follows

$$\mathcal{I} = \int_\tau^0 \frac{e^{Ms} - 1}{1 - Ne^{Ms}} = -\int_\tau^0 \frac{e^{Ms} - 1}{Ne^{Ms} - 1} ds \tag{264}$$

We begin by making the substitution $x = -Ms \implies dx = -Mds$:

$$\mathcal{I} = \frac{1}{M} \int_{-M\tau}^0 \frac{e^{-x} - 1}{Ne^{-x} - 1} dx \tag{265}$$

$$= \frac{1}{M} \int_{-M\tau}^0 \frac{e^x - 1}{e^x - N} dx \tag{266}$$

$$= \frac{1}{M} \int_{-M\tau}^0 \frac{e^x - N}{e^x - N} + \frac{N - 1}{e^x - N} dx \tag{267}$$

$$= \frac{1}{M} x\Big|_{-M\tau}^0 + \frac{1}{M} \int_{-M\tau}^0 \frac{N - 1}{e^x - N} dx \tag{268}$$

$$= \tau + \frac{N - 1}{M} \int_{-M\tau}^0 \frac{1}{e^x - N} dx \tag{269}$$

We now let $y = e^x \implies \frac{dy}{y} = dx$

$$\mathcal{I} = \tau + \frac{N - 1}{M} \int_{e^{-M\tau}}^1 \frac{1}{y(y - N)} dy \tag{270}$$

$$\mathcal{I} = \tau + \frac{N - 1}{M} \int_{e^{-M\tau}}^1 \frac{1}{y^2(1 - \frac{N}{y})} dy \tag{271}$$

$$\mathcal{I} = \tau + \frac{N - 1}{MN} \log\left(1 - \frac{N}{y}\right)\Bigg|_{e^{-M\tau}}^1 \tag{272}$$

$$\mathcal{I} = \tau + \frac{N - 1}{MN}\left[\log(1 - N) - \log(1 - Ne^{M\tau})\right] \tag{273}$$

$$\mathcal{I} = \tau + \frac{N - 1}{MN} \log\left(\frac{1 - N}{1 - Ne^{M\tau}}\right) \tag{274}$$

So we can return back to $A$ as:

$$A(\tau,\varphi) = i\varphi r\tau + \frac{\kappa\theta}{\xi^2}(i\varphi\rho\xi + b_j - M)\left[\tau + \frac{N-1}{MN}\log\left(\frac{1-N}{1-Ne^{M\tau}}\right)\right] \tag{275}$$

$$= i\varphi r\tau + \frac{\kappa\theta}{\xi^2}\left[(i\varphi\rho\xi + b_j - M)\tau + (i\varphi\rho\xi + b_j - M)\frac{N-1}{MN}\log\left(\frac{1-N}{1-Ne^{M\tau}}\right)\right] \tag{276}$$

We now consider the logarithm pre-factor:

$$(i\varphi\rho\xi + b_j - M)\frac{N-1}{MN} = (i\varphi\rho\xi + b_j - M)\frac{\frac{i\varphi\rho\xi+b_j-M}{i\varphi\rho\xi+b_j+M} - 1}{M\frac{i\varphi\rho\xi+b_j-M}{i\varphi\rho\xi+b_j+M}} \tag{277}$$

$$= (i\varphi\rho\xi + b_j + M)\frac{\frac{i\varphi\rho\xi+b_j-M}{i\varphi\rho\xi+b_j+M} - 1}{M} \tag{278}$$

$$= \frac{(i\varphi\rho\xi + b_j - M) - (i\varphi\rho\xi + b_j + M)}{M} \tag{279}$$

$$= -\frac{2M}{M} \tag{280}$$

$$= -2 \tag{281}$$

Hence we see that

$$A(\tau,\varphi) = i\varphi r\tau + \frac{\kappa\theta}{\xi^2}\left[(i\varphi\rho\xi + b_j - M)\tau - 2\log\left(\frac{1-N}{1-Ne^{M\tau}}\right)\right] \tag{282}$$

# F Poisson Process

In this appendix we give a brief overview of the theoretical structure of the Poisson process. In its simplest description, a Poisson process is a counting process that tracks the number of random events that have occurred before some (deterministic) time $t$ ie

**Definition F.1** (Poisson Process)**.** *A (homogenous) Poisson process, $N(t)$ is a counting process satisfying the following conditions:*

1. $N(0) = 0$ *and for* $s < t$ *then* $N(s) \leqslant N(t)$

2. $\mathbb{P}\left[N(t+h) = n+m | N(t) = n\right] = \begin{cases} 1 - \lambda h + o(h) & m = 0 \\ \lambda h + o(h) & m = 1 \\ o(h) & m > 1 \end{cases}$

3. *if* $s < t$*, then* $N(t) - N(s)$ *is independent of* $N(s)$

Note in particular, the second condition which states the in each time period $(t, t+h]$, the probability that the Poisson process increases by 1 can be viewed as Bernoulli trial. In conjunction with the third condition, this gives a way of teasing out the distribution of the Poisson process by making use of a Binomial approximation by making use of a Poisson Limit Theorem style argument (we will go into further detail of this in the next section as a motivation for the one of the potential constructions for the Poisson process). For now, we state and prove the form of the distribution of the Poisson process

**Theorem F.2.** *$N(t)$ has a Poisson($\lambda t$) distribution, that is*

$$\mathbb{P}\left[N(t) = n\right] = \frac{(\lambda t)^n e^{-\lambda t}}{n!} \tag{283}$$

*Proof.* The proof will take the following course. First, we will derive a condition that relates the derivatives of the probability to itself. Second, we shall make use of generating functions to give a full distribution.
For clarity of expression, for the rest of this proof, we let $p_n(t) = \mathbb{P}[N(t) = n]$

$$\begin{aligned}
p_n(t+h) = \mathbb{P}[N(t+h) = n] &= \sum_m \mathbb{P}[N(t+h) = n | N(t) = m]\mathbb{P}[N(t) = m] && (284)\\
&= (1 - \lambda h + o(h))\mathbb{P}[N(t) = n] && (285)\\
&\quad + (\lambda h + o(h))\mathbb{P}[N(t) = n-1] && (286)\\
&\quad + o(h)\mathbb{P}[N(t) \leqslant n-2] && (287)\\
&= (1 - \lambda h)p_n(t) + \lambda h p_{n-1}(t) + o(h) && (288)\\
\implies \quad \frac{p_n(t+h) - p_n(t)}{h} &= \lambda p_{n-1}(t) - \lambda p_n(t) + o(1) && (289)\\
\implies p_n'(t) &= \lambda(p_{n-1}(t) - p_n(t)) && (290)
\end{aligned}$$

In conjunction with an initial condition that $p_0'(t) = -\lambda p_0(t)$, we have a system that we can solve for, concluding the first part of this proof

To finish this proof, let

$$G(s,t) = \sum_{n=0}^{\infty} p_n(t)s^n = \mathbb{E}\left[s^{N(t)}\right] \tag{291}$$

We now consider the derivative of the $G$ with respect to $t$:

$$\begin{aligned}
\frac{\partial}{\partial t}G(s,t) = \sum_{n=0}^{\infty} p_n'(t)s^n &= \sum_{n=0}^{\infty} \lambda(p_{n-1}(t) - p_n(t))s^n && (292)\\
&= \lambda\sum_{n=0}^{\infty} p_{n-1}(t)s^n - \lambda\sum_{n=0}^{\infty} p_n(t)s^n && (293)\\
&= \lambda s G(s,t) - \lambda G(s,t) && (294)
\end{aligned}$$

Hence we have to solce

$$\frac{\partial}{\partial t}G(s,t) = \lambda(s-1)G(s,t) \qquad G(s,0) = 1 \tag{295}$$

which has solution

$$G(s,t) = e^{-(1-s)\lambda t} = \sum_{n=0}^{\infty}\left(\frac{(\lambda t)^n e^{-\lambda t}}{n!}\right)s^n \tag{296}$$

hence we see that $\mathbb{P}[N(t) = n] = \frac{(\lambda t)e^{-\lambda t}}{n!}$ or rather, $N(t)$ has a Poisson($\lambda t$) distribution ☐

# References


Black, F. & Scholes, M. (1972), 'The valuation of option contracts and a test of market efficiency', *The journal of finance* **27**(2), 399–417.

Brummelhus, R. (2018), 'Solving pdes using probabiity: The feynman-kac method'.
**URL:** *http://econ109.econ.bbk.ac.uk/brad/Methods/FeynmanKac/Feynman%20-%20Kac%20lecture.pdf*

Carr, P. & Madan, D. (1999), 'Option valuation using the fast fourier transform', *Journal of computational finance* **2**(4), 61–73.

Cox, J. C., Ingersoll, J. E. & Ross, S. A. (1985), 'A theory of the term structure of interest rates', *Econometrica* **53**(2), 385–407.
**URL:** *http://www.jstor.org/stable/1911242*

Dunn, R., Hauser, P., Seibold, T. & Gong, H. (2014), 'Estimating option prices with heston's stochastic volatility model'.

Feller, W. (1951), 'Two singular diffusion problems', *Annals of Mathematics* **54**(1), 173–182.
**URL:** *http://www.jstor.org/stable/1969318*

Gatheral, J. (2011), *The volatility surface: a practitioner's guide*, John Wiley & Sons.

Gikhman, I. I. (2011), 'A short remark on feller's square root condition', *Available at SSRN 1756450* .

Gil-Pelaez, J. (1951), 'Note on the inversion theorem', *Biometrika* **38**(3-4), 481–482.

Heston, S. L. (1993), 'A closed-form solution for options with stochastic volatility with applications to bond and currency options', *The review of financial studies* **6**(2), 327–343.

Kahl, C. & Jäckel, P. (2005), 'Not-so-complex logarithms in the heston model', *Wilmott magazine* **19**(9), 94–103.

Lalley, S. (2001), 'Stochastic differential equations, diffusion processes, and the feynman-kac formula'.
**URL:** *http://www.stat.uchicago.edu/ lalley/Courses/391/Lecture12.pdf*

Lee, R. W. et al. (2004), 'Option pricing by transform methods: extensions, unification and error control', *Journal of Computational Finance* **7**(3), 51–86.

Matsuda, K. (2004), 'Introduction to merton jump diffusion model', *Department of Economics, The Graduate Center, The City University of New York, New York* .

Merton, R. C. (1976), 'Option pricing when underlying stock returns are discontinuous', *Journal of financial economics* **3**(1-2), 125–144.

Rouah, F. D. (2013), 'Simplified derivation of the heston model', *Fabrice Rouah* .

Rouah, F. D. (2020), 'Four derivations of the black scholes pde', *Mathematical finance notes. http://www. frouah. com/pages/finmath. html. Accessed July* .

Schmelzle, M. (2010), 'Option pricing formulae using fourier transform: Theory and application', *Preprint, http://pfadintegral. com* .

Schöbel, R. & Zhu, J. (1999), 'Stochastic volatility with an ornstein–uhlenbeck process: an extension', *Review of Finance* **3**(1), 23–46.

Sepp, A. (2003), 'Pricing european-style options under jump diffusion processes with stochastic volatility: Applications of fourier transform', *Kangro, R., Parna, K., and Sepp, A.,(2004)," Pricing European-Style Options under Jump Diffusion Processes with Stochastic Volatility: Applications of Fourier Transform", Acta et Commentationes Universitatis Tartuensis de Mathematica* **8**, 123–133.

Tankov, P. (2003), *Financial modelling with jump processes*, Chapman and Hall/CRC.

Ye, Z. (2013), The black-scholes and heston models for option pricing, Master's thesis, University of Waterloo.