

MATHEMATICAL MODELING OF DERIVATIVE PRICING

WILL GRIFFIN

ABSTRACT. We introduce mathematical modeling and related formulas for option pricing through two key European options. In the first part, we derive the Black-Scholes-Merton PDE and solution. In the second part, we discuss its shortcomings and turn to stochastic volatility modeling, in particular the Heston Model.

CONTENTS

1. Introduction	1
2. Review of Stochastic Calculus Results	2
3. The Black-Scholes-Merton Model	3
3.1. Market Model	3
3.2. PDE Derivation	5
3.3. Discussion of Assumptions	8
4. Stochastic Volatility	9
4.1. Observing the Volatility Parameter	9
4.2. The Stochastic Volatility Paradigm	10
4.3. The Heston Model	11
4.4. Solving the Heston PDE	14
5. Appendix: Girsanov Theorem and Feynman-Kac	15
Acknowledgments	17
References	17

1. INTRODUCTION

The goal of this paper is to introduce a reader with some familiarity with stochastic processes to intermediate option pricing. One of the standard models in this field is the Black-Scholes-Merton model. Since it was first developed in the 1970's, the question of how to price options has been one of the most central questions in mathematical finance. One will often first hear of Black-Scholes-Merton as the solution to this question, and it will continue to be alluded to until one develops enough mathematical maturity to grasp stochastic calculus. However, it has been apparent for some time that the model makes several critically incorrect assumptions, and it is seldom the go-to model of traders today. Why, then, is it still so ubiquitous? Part of the reason is likely its simplicity. Black-Scholes-Merton is an incredibly useful tool to introduce oneself to options pricing arguments, such as leveraging the assumption of no-arbitrage. This paper does exactly that, but elects

Date: August 29, 2022.

to spend less time with the subtleties of stochastic calculus in favor of discussing stochastic volatility models at greater length.

2. REVIEW OF STOCHASTIC CALCULUS RESULTS

This paper introduces the reader to some intermediate concepts in mathematical finance. This requires familiarity with the basics of stochastic calculus, which are stated here for completeness, but are not proven at length. For a detailed treatment, the reader is encouraged to consult [2]. Notice that a few measure-theoretic definitions are stated, but intimate knowledge of measure theory is not a hard requirement for the interested reader. These statements and results are drawn from [3].

Definition 2.1. (Brownian Motion)

An **n-dimensional Brownian Motion**, denoted \vec{W}_t , adapted to a filtered probability space¹ $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t), \mathbb{P})$ ² is an \mathbb{R}^n -valued stochastic process satisfying:

- $\vec{W}_0 = \vec{0}$;
- (Independent Increments) $(\vec{W}_t - \vec{W}_s)$ is independent of $\mathcal{F}_s \forall s \leq t$;
- (Stationarity)³ $(\vec{W}_t - \vec{W}_s) \sim N_d(0, (t - s)I_d) \forall s \leq t$;
- (Continuous Sample Paths) The sample path $\phi : t \rightarrow \vec{W}_t$ is continuous \mathbb{P} -almost surely;

Definition 2.2. (Martingales)⁴

A **Martingale** M_t is a stochastic process satisfying:

- $\mathbb{E}[|M_t|] < \infty \forall t \in \mathbb{R}$
- $\mathbb{E}[M_t | M_s] = M_s \forall s \leq t$

Theorem 2.3. *An n-dimensional Brownian Motion \vec{W}_t is a Martingale.*

Definition 2.4. (Quadratic Covariation)

The **Quadratic Covariation** of two Stochastic Processes X_t and Y_t is given by:

$$[X, Y]_t = \lim_{\|P\| \rightarrow 0} \sum_{k=1}^n (X_{t_k} - X_{t_{k-1}})(Y_{t_k} - Y_{t_{k-1}})$$

Here, $\|P\|$ represents the mesh size of the regular partition $\{t_1 \dots t_n\}$. The **Quadratic Variation** of a process is given by $[X, X]_t$, sometimes denoted $[X]_t$

Theorem 2.5. (Important Covariations)⁵

Let W_t be a standard Brownian Motion. Then we have:

- $[W, t]_t = 0$
- $[t, t]_t = 0$
- $[W, W]_t = t$

¹For our purposes, the filtration is assumed to be the standard filtration $\mathcal{F}_t = \sigma(W_t)$.

²All stochastic processes we describe from here on out are assumed to be adapted to a filtered probability space in a similar way.

³Here I_d denotes the Identity Matrix in n dimensions.

⁴ $\mathbb{E}[M_t | M_s]$ represents conditional expectation.

⁵Here, t is just the current time, a nonrandom process.

Convention 2.1. (Box Calculus)

The following convention is used when dealing with Itô Calculus. While it is non-rigorous⁶, it is a useful heuristic that is computationally correct for our purposes.

- $d[X^i, X^j]_t = dX_t^i dX_t^j$
- $dW_t dt = 0$
- $dt dt = 0$
- $dW_t dW_t = dt$

Theorem 2.6. (Multivariate Itô's Formula)⁷

Let f be a twice continuously differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and \vec{X}_t be a continuous semimartingale in \mathbb{R}^n ⁸. Then we have:

$$(2.7) \quad f(\vec{X}_t) = f(\vec{X}_0) + \sum_{i=1}^n \int_0^t \partial_{x_i} f(\vec{X}_s) dX_s^i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \int_0^t \partial_{x_i x_j} f(\vec{X}_s) d[X^i, X^j]_s$$

In Differential Notation:

$$(2.8) \quad df(\vec{X}_t) = \sum_{i=1}^n \partial_{x_i} f(\vec{X}_t) dX_t^i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \partial_{x_i x_j} f(\vec{X}_t) d[X^i, X^j]_t$$

3. THE BLACK-SCHOLES-MERTON MODEL

3.1. Market Model. In this section we will introduce the market model for option pricing, then derive the Black-Scholes-Merton PDE and pricing formula. Black-Scholes-Merton, despite being outdated, provides a perfect example of the mathematical reasoning behind pricing of financial assets. We will then discuss the construction at greater length and motivate why stochastic volatility is a natural step forward beyond the classical model. We will not assume the reader has any background in finance and build the model up from scratch, starting with the definition of a European Option.

Definition 3.1. (European Options)

A **European Option** on an asset in the market, called the **underlying asset/underlying**, is a contract that grants the owner the right, but not the obligation to buy (call contract) or sell (put contract) the underlying for a predetermined price at a predetermined date. The price of the underlying is modeled by a stochastic process denoted X_t . The option contract is characterised by a **strike price**, the price at which the user is entitled to buy/sell, denoted κ , and an **expiry date/maturity**, the time at which the user is entitled to exercise the rights granted by the option, denoted τ . At maturity, the option contract will grant a payoff of

$$(X_\tau - \kappa)^+ = \max\{(X_\tau - \kappa), 0\}$$

The main question we want to answer is "How much should it cost to purchase a European option contract?" For now, assume the option is a call. If we purchase the option (often called **longing** the option) we want the underlying asset price

⁶Algebraic operators beyond the scope of this paper can make this rigorous

⁷Note we use $\partial_{x_i} f$ to denote $\frac{\partial f}{\partial x_i}$

⁸This is a technical requirement that is assumed about any process on which we invoke Itô's Lemma in this paper

to go up past the strike price so that we can exercise the option, buy it for lower than the market price, and then be able to sell immediately for profit⁹. If we sell the option, we incur the risk of the option being exercised against us (often called **shorting** the option), and we do not want the owner to do so. Therefore, we want the underlying to stay low enough so that the owner does not choose to exercise it and profit. To tackle the problem of finding a fair price, we need to make modeling assumptions about how the markets work, and how different asset prices evolve over time.

Assumption 3.1. (Market Assets)

Our wealth over time is given by our **portfolio value**, an evolving process denoted π_t , which sums the value of all our owned securities at each time t . We can trade two non-derivative assets, a **risky asset/stock** and a **riskless asset/bond**. The values of these assets through time are characterized by the following processes:

- (Bond) $dB_t = rB_t dt \Rightarrow B_t = B_0 e^{rt}$
- (Stock) $dX_t = \mu X_t dt + \sigma X_t dW_t \Rightarrow X_t = X_0 \exp\{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t\}$

Notice that the bond value is a deterministic (nonrandom) process, and the Stock Price is a stochastic process called **Geometric Brownian Motion**. μ represents mean returns of the stock, σ represents volatility of the stock, and r represents the continuously compounding rate of interest on the bond. Assuming these are constant through time is a very strong statement that we are going to revisit, as it is the primary weakness of Geometric Brownian Motion as a pricing model for stocks.

Assumption 3.2. (Short Selling)

We are allowed to short any security. Shorting an asset (not a derivative), is when we sell something we do not own and incur owing that value back at some point. It appears as a negative value when calculating our total portfolio value¹⁰. A **security** is any financial instrument that is traded in the market, so all stocks, bonds, and derivatives in the market.

Assumption 3.3. (Self-Financing Portfolio)

When we begin trading securities at $t = 0$, we will have an initial amount of money to make purchases (often it is assumed to be 0). We cannot invest any additional money beyond this amount as time progresses. If we want more money with which to trade, we are going to get it by shorting securities (options or assets) and reinvesting that cash into other securities. Our portfolio is said to be **Self-Financing** when we generate money to trade this way. Formally, let π_t be our wealth over time with π_0 our initial investment. Suppose we can trade N assets whose prices through time are denoted X_t^i , with A_t^i denoting the number of shares we own of the i th asset. Our assumption is then formally stated:

$$\pi_t = \pi_0 + \sum_{i=1}^N \int_0^t A_t^i dX_t^i$$

⁹Assuming that the gain from exercising is big enough to offset the cost of buying the option.

¹⁰Shorting is quite complex in practice and covers a wide range of methods to bet against a security. However, for this paper, knowing what being short an asset and a European derivative is sufficient.

In differential form:

$$(3.2) \quad d\pi_t = \sum_{i=1}^N A_t^i dX_t^i$$

Assumption 3.4. (No Arbitrage)

There are no arbitrage opportunities in the market. **Arbitrage** is when an actor in the market can make money risk-free from nothing due to inefficiencies in pricing all the securities.

Assumption 3.5. (No Dividends)

Stocks do not pay cash dividends to owners. A **dividend** is a cash portion of company profits issued to shareholders as partial owners of the company.

Assumption 3.6. (No Transaction Fees)

Market actors do not incur any fees to make trades.

Assumption 3.7. (Divisible Securities)

All securities are perfectly divisible, so any dollar amount can be invested into any security.

Assumption 3.8. (Complete Markets)

Assuming complete markets means we can **replicate** any position (long or short) on any security in the market. To replicate a position in a security, one trades other existing market assets to create a portfolio that exactly imitates the security payoff. Buying this **replicating portfolio** is exactly the same as buying the security, which means they should cost the same to buy, otherwise there would be arbitrage. This principle is often called the **Law of One Price**. Formally, if X_t is the price of a security in the market and π_t is our portfolio, we are assured that it is possible to trade assets to induce the following condition:

$$d\pi_t = dX_t$$

3.2. PDE Derivation. Consider a market satisfying the assumptions we have outlined. The option contract is characterized by two things: expiry τ and strike price κ , so the price should be a function of the underlying and time. Denote this function of time and underlying asset price $P(t, x)$. Assume $\pi_0 = 0$, i.e. we do not start with any money to trade. Because our portfolio is self-financing, we must short some security to generate capital that we can trade. Over time, we will short the option contract. Denote units shorted over time N_t . We will then invest the capital into the risky stock. Denote units held over time A_t . Thus our portfolio value evolves by the following specific case of (3.2):

$$(3.3) \quad d\pi_t = A_t dX_t - N_t dP(t, X_t)$$

Right now the only securities in our market are the derivative, risky stock, and riskless bond. Because markets are complete, we can replicate the riskless asset by trading the derivative and risky asset. Thus, we can assume that our trading positions A_t and N_t evolve over time to replicate and impose the formal condition:

$$(3.4) \quad d\pi_t = r\pi_t dt$$

Plugging in (3.3), this expands to:

$$(3.5) \quad A_t dX_t - N_t dP(t, X_t) = r(A_t X_t - N_t P(t, X_t)) dt$$

We do not yet know what $dP(t, X_t)$ is, but we can apply Itô's formula to find out:

$$\begin{aligned}
dP(t, X_t) &= \partial_t P(t, X_t)dt + \partial_x P(t, X_t)dX_t + \frac{1}{2}\partial_{xx}P(t, X_t)dX_t dX_t \\
&= \partial_t P(t, X_t)dt + \partial_x P(t, X_t)(\mu X_t dt + \sigma X_t dW_t) + \frac{1}{2}\partial_{xx}P(t, X_t)(\sigma^2 X_t^2 dW_t dW_t) \\
(3.6) \quad &= (\partial_t P(t, X_t) + \mu X_t \partial_x P(t, X_t) + \frac{1}{2}\sigma^2 X_t^2 \partial_{xx}P(t, X_t))dt + (\sigma X_t \partial_x P(t, X_t))dW_t
\end{aligned}$$

Plugging (3.6) into (3.5), we obtain:

$$\begin{aligned}
&A_t(\mu X_t dt + \sigma X_t dW_t) \\
-N_t\{[\partial_t P(t, X_t) + \mu X_t \partial_x P(t, X_t) + \frac{1}{2}\sigma^2 X_t^2 \partial_{xx}P(t, X_t)]dt + [\sigma X_t (\partial_x P(t, X_t))]dW_t\} \\
&= r(A_t X_t - N_t P(t, X_t))dt
\end{aligned}$$

The right hand side of this equation has no stochastic term (or $0dW_t$ if you prefer), so we set the left hand side stochastic term to 0:

$$\sigma A_t X_t - \sigma \partial_x P(t, X_t) N_t X_t = 0$$

$$(3.7) \quad \Rightarrow A_t = N_t \partial_x P(t, X_t)^{11}$$

Equating the dt terms gives us:

$$\begin{aligned}
\mu A_t X_t dt - N_t \partial_t P(t, X_t) dt - \mu N_t \partial_x P(t, X_t) X_t dt \\
- \frac{1}{2}\sigma^2 N_t X_t^2 \partial_{xx}P(t, X_t) dt = r(A_t X_t - N_t P(t, X_t)) dt
\end{aligned}$$

Using (3.7) to substitute away any instances of A_t we get:

$$\Rightarrow -N_t[\partial_t P(t, X_t) + \frac{1}{2}\sigma^2 X_t^2 \partial_{xx}P(t, X_t)]dt = r(N_t X_t \partial_x P(t, X_t) - N_t P(t, X_t))dt$$

Cancelling the N_t terms and moving everything to one side leaves us with the **Black-Scholes PDE**:

$$(3.8) \quad \partial_t P(t, X_t) + \frac{1}{2}\sigma^2 X_t^2 \partial_{xx}P(t, X_t) + r(X_t \partial_x P(t, X_t) - P(t, X_t)) = 0$$

Recall that once the option matures at time τ , its price **must** be the same as the payoff from exercising it by the Law of One Price. This is our terminal condition:

$$(3.9) \quad P(\tau, x) = \max\{x - \kappa, 0\}$$

The existence/uniqueness of a solution to this PDE follows by existence/uniqueness of a solution to the Heat Equation, which is treated in detail by [16].

Notice that P is a function with a time and "space" (price) variable, and our input in the space variable is a random. However, we can solve the PDE with deterministic inputs and simply plug in the observed state of X_t to the deterministic solution.

¹¹Note X_t is never zero by construction

Theorem 3.10.

$$(3.11) \quad P(t, X_t) = X_t \Phi(d_1) - \kappa e^{-(\tau-t)} \Phi(d_2)$$

where Φ represents the CDF of a standard normal distribution and:

$$d_1 = \frac{\log\left(\frac{X_t}{\kappa} + \left(r + \frac{1}{2}\sigma^2\right)(\tau - t)\right)}{\sigma\sqrt{\tau - t}}$$

$$d_2 = d_1 - \sigma\sqrt{\tau - t}$$

This result can be proven with probability using risk-neutral probability measures from measure theory or through differential equation methods related to the Heat Equation. The latter is presented here, adapted from [4] to match our notation thus far:

Proof. We have the Black Scholes PDE from (3.8) under the terminal condition that the option is worth its payoff from exercising when it hits maturity:

$$(3.12) \quad P(\tau, x) = \max\{x - \kappa, 0\}$$

We also impose the boundary condition that reflects that an option is worthless if the underlying is worthless ¹²:

$$P(t, 0) = 0 \quad \forall t \in [0, \tau]$$

However, by construction of X_t , any stock input will satisfy this. Now, execute the following change of variables from the (t, x) domain to a (u, v) domain:

$$x = e^v$$

$$t = \tau - \frac{2u}{\sigma^2}$$

Thus we get:

$$P(t, x) = p(u, v) = p\left(\frac{\sigma^2(\tau - t)}{2}, \ln(x)\right)$$

By the Chain Rule, the partials of P in terms of the (u, v) variables become:

$$\partial_t P(t, x) = \frac{-\sigma^2}{2} \partial_u p(u, v)$$

$$\partial_x P(t, x) = e^{-v} \partial_v p(u, v)$$

$$\partial_{xx} P(t, x) = -(e^{-v})^2 \partial_v^2 p(u, v) + (e^{-v})^2 \partial_{vv} p(u, v)$$

Substituting these into the Black-Scholes PDE (3.8), and simplifying we get:

$$(3.13) \quad \partial_u p(u, v) = \partial_{vv} p(u, v) + \left(\frac{2r}{\sigma^2} - 1\right) \partial_v p(u, v) - \frac{2r}{\sigma^2} p(u, v)$$

Store the parameter expression $\frac{2r}{\sigma^2}$ in a single parameter k and get:

$$(3.14) \quad \partial_u p(u, v) = \partial_{vv} p(u, v) + (k - 1) \partial_v p(u, v) - k p(u, v)$$

Now the transformed problem is (3.12) subject to boundary conditions:

$$(3.15) \quad p(0, v) = P(\tau, e^v) = \max\{e^v - \kappa, 0\}$$

Notice that evaluating p at 0 in transformed time domain u is equivalent to evaluating P at expiry in the original domain. Thus, output of p has to obey the same constraint using the transformed space domain. We will execute another

¹²This will end up irrelevant since Geometric Brownian Motion is never 0

change of variables to make our constrained PDE resemble the Heat Equation. Let $\phi(u, v) = e^{\alpha v + \beta u}$ where α and β are parameters to be determined later and set:

$$(3.16) \quad p(u, v) = \phi(u, v)q(u, v)$$

Computing the partials, we get:

$$\begin{aligned} \partial_u p(u, v) &= \beta \phi(u, v)q(u, v) + \phi(u, v)\partial_u q(u, v) \\ \partial_v p(u, v) &= \alpha \phi(u, v)q(u, v) + \phi(u, v)\partial_v q(u, v) \\ \partial_{vv} p(u, v) &= \alpha^2 \phi(u, v)q(u, v) + 2\alpha \phi(u, v)\partial_v q(u, v) + \phi(u, v)\partial_{vv} q(u, v) \end{aligned}$$

Now choose the following convenient parameters:

$$\begin{aligned} \alpha &= -\frac{1}{2}(k-1) = \frac{\sigma^2 - 2r}{2\sigma^2} \\ \beta &= -\frac{1}{4}(k+1)^2 = -\left(\frac{\sigma^2 + 2r}{2\sigma^2}\right)^2 \end{aligned}$$

Substituting the partials and parameters into (3.14) we obtain:

$$(3.17) \quad \partial_u q(u, v) = \partial_{vv} q(u, v); \quad |v| < \infty, \quad u \in [0, \frac{\sigma^2 \tau}{2}]$$

Our transformed boundary condition is now:

$$(3.18) \quad q(0, v) = e^{-\alpha v} p(0, v) = e^{-\alpha x} \max\{e^x - \kappa, 0\}$$

As it turns out this differential equation and boundary condition are the well-known heat equation [13], which admits the following solution:

$$(3.19) \quad q(u, v) = \frac{1}{\sqrt{4\pi u}} \int_{\mathbb{R}} q(0, z) e^{-\frac{(v-z)^2}{4u}} dz$$

Transforming back to our original domains then yields the result of the theorem. \square

3.3. Discussion of Assumptions. Statistician George E.P. Box once famously said "All models are wrong, but some are useful." Economic models are no different; most assumptions are wrong in some sense, but the value of a model, wrong assumptions or not, only starts to wane where its predictive power starts to fail. The goal of this section is to review where Black-Scholes assumptions are wrong (almost everywhere), and to discuss when and where these assumptions actually begin to harm the model.

Objection 3.1. Arbitrage does exist.

This might be the most obviously wrong assumption of the model, but is interestingly not very harmful over short periods of time. The advent of automated, high frequency trading has led to minuscule arbitrage opportunities quickly being grabbed and corrected by extremely high performance machines faster than a human can think, and humans can often ignore those arbitrage situations as a result, allowing computers to handle it. A more detailed discussion can be found here [8].

Objection 3.2. Transactions are not free.

This can change a trading strategy as transaction costs can move a narrowly profitable strategy into unprofitable territory, accounting for having to pay the transaction costs associated with acquiring or selling the necessary securities. However, this problem has at least been fixed by newer versions of the model that do account for transaction costs, e.g. [9].

Objection 3.3. Plenty of stocks pay dividends.

This also changes the profitability of trading a risk-free portfolio and invalidates a key underlying piece of the mathematical conclusions we came to. However, this has also since been studied and an introduction can be found here [10].

Objection 3.4. Shorting and continuous time trading are not guaranteed.

For larger firms making trades, there are often administrative lags for quickly turning over lots of capital into the markets (regulations, infrastructure, etc), so they cannot trade continuously. Nor is every security available to be shorted. However, technology allows trading times to be "approximately continuous" and enough securities can be shorted such that this is more of an annoyance than a model-breaking catastrophe. One merely needs to ensure their analysis is restricted to markets of "shortable" assets.

Objection 3.5. Bond Markets are oversimplified.

Bond markets are highly complex and "risk free" rates of return do not really exist. In this context, risk refers to the chance that the issuer defaults on their debt. The US government, which issues US treasury bonds, is considered the best approximation of a risk free bond writer, as the probability of the US collapsing is considered negligibly low. However, there are so many possible bond writers that exhibit extremely low risk of default at varying rates of return; almost every government and firm issues debt that may or may not be available to purchase. This cannot be encapsulated by a single parameter.

Objection 3.6. Stocks do not follow Geometric Brownian Motion

Geometric Brownian Motion is far too "thin tailed"¹³ to accurately model the potentially extreme movements of stocks, particularly periodic crashes. An extreme example would be the 1987 stock market crash on Black Monday, in which the Dow Jones Industrial Average collapsed. The probability of this event was approximately $\frac{1}{10^{50}}$ [11]. However, stochastic volatility, the subject of the second half of this paper, can go a long way in addressing this concern, which is arguably the biggest problem with Black-Scholes Merton.

Objection 3.7. Markets are not Complete

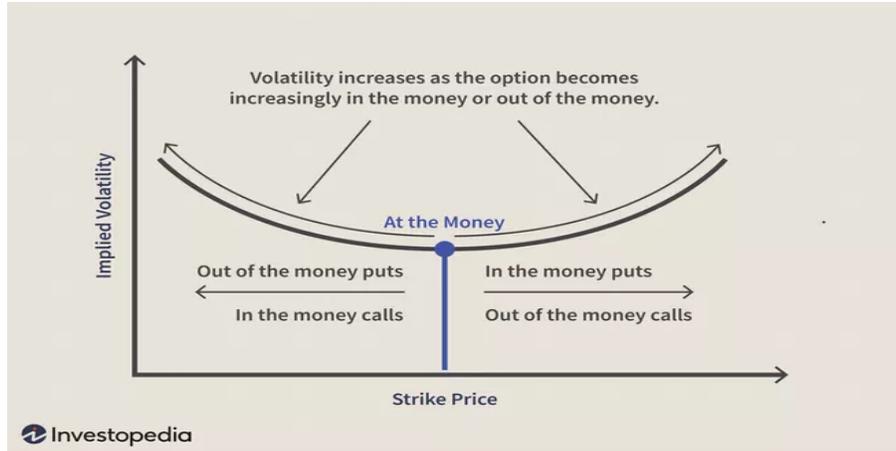
Complete Markets are intimately tied to the Fundamental Theorems of Asset Pricing. These are deep results that rely on measure theoretic analysis of risk neutral probability measures, which would hinder the accessibility of this paper. Briefly, more developed markets, such as the NYSE, can approximate completeness, but less developed markets, such as a stock exchange in a nation that has recently undergone catastrophic civil conflict, will break this assumption more often.

4. STOCHASTIC VOLATILITY

4.1. Observing the Volatility Parameter. With the Black-Scholes-Merton formula in hand, one can derive a theoretical price for an option based on the current state of the random stock process. However, in practice, multiple quotes for the same option (same strike price and expiry) can be discovered. How is this possible? The volatility parameter, σ of the stock movement X_t is **not** directly observable, meaning individuals who trade in the market have to estimate what exactly this parameter is. In practice, traders come up with very different opinions

¹³There is very little probability density around tails of the distribution

about that particular value. This opinion is always observable (when pricing with Black-Scholes-Merton) upon observing the price of the option contract, we can manipulate the explicit pricing formula to solve for the **implied volatility** (the value of σ that the trader must be using to offer the price) contained in the offer. As it turns out, plotting implied volatility of an option as a function of its strike price (holding expiry constant) leads to a significant level of curvature. This phenomenon is known as the **volatility smile** [12].



This immediately tells us that σ is not constant through time, and multiple models of the non-constant behavior of volatility have been proposed since the discovery of this phenomenon.

4.2. The Stochastic Volatility Paradigm. Recall the classical SDE designed to describe stock behavior:

$$dX_t = \mu X_t dt + \sigma X_t dW_t^0$$

The Stochastic Volatility model adjusts to the following SDE:

$$dX_t = \mu_t X_t dt + \sigma_t X_t dW_t^0$$

μ_t and σ_t are now also stochastic processes evolving randomly through time. They are called the **returns process** and **volatility process**. Generally, the volatility process is modeled as a function¹⁴ of a **driving process**, denoted Y_t and satisfying the SDE:

$$dY_t = \alpha(Y_t)dt + \beta(Y_t)dW_t^1; \beta(Y_t) \geq 0$$

So then

$$\sigma_t = f(Y_t)$$

Here we should be very careful, as there are now two Brownian Motions evolving through time in our market. W_t^0 is the Brownian Motion driving the stock's random movements; W_t^1 drives the volatility's random movements. Further, these Brownian motions might **not** be independent. They have a correlation coefficient denoted ρ . More precisely:

$$[dW_t^0, dW_t^1] = \rho; \rho \in [-1, 1]$$

How the driving process behaves is another modeling debate. For instance, it is desirable that the driving process pulls back to a somewhat regular value, as this

¹⁴Assumed to be well behaved enough to apply Itô's Formula

fits actual observations about volatility data [14]. This can be accomplished by modeling volatility with a **mean reverting** driving process, which looks like this:

$$dY_t = \alpha(m - Y_t)dt + \beta(Y_t)dW_t^1$$

Here, m is the **mean** and α is the **rate of mean reversion**. Notice that whenever $Y_t > m$, the drift term is negative (downward drift towards m) and whenever $Y_t < m$, the drift term is positive (upward drift towards m). The strength of the drift will go up the further Y_t strays from m , leading to a "rubber band" effect on Y_t , pulling it towards m . [1] provides several examples of common driving processes in the literature:

$$dY_t = \alpha Y_t dt + \beta Y_t dW_t^1 \quad (\text{Logonormal})$$

$$dY_t = \alpha(m - Y_t)dt + \beta W_t^1 \quad (\text{Ornstein - Uhlenbeck})$$

$$dY_t = \alpha(m - Y_t)dt + \beta\sqrt{Y_t}dW_t^1 \quad (\text{Cox - Ingrell - Ross})$$

[1] also summarizes several popular stochastic volatility models that can be built on these tools so far:

- **Hull-White:** Logonormal driver, $\rho = 0$, $\sigma_t = \sqrt{Y_t}$
- **Scott:** Ornstein-Uhlenbeck driver, $\rho = 0$, $\sigma_t = e^{Y_t}$
- **Stein-Stein:** Ornstein-Uhlenbeck driver, $\rho = 0$, $\sigma_t = |Y_t|$
- **Ball-Roma:** Cox-Ingrell-Ross driver, $\rho = 0$, $\sigma_t = \sqrt{Y_t}$
- **Heston:** Cox-Ingrell-Ross driver, $|\rho| > 0$, $\sigma_t = \sqrt{Y_t}$

We will focus on the Heston model for the remainder of the paper, as it is generally the most ubiquitous and, critically, recognizes the correlation between the Brownian Motion driving the stock price and the Brownian Motion driving volatility, which is a well-documented phenomenon [15].

4.3. The Heston Model. First, we restate all the assumptions the Heston makes for clarity. Notice we are taking the square root of σ_t in the stock process:

- (Stock Dynamics) $dX_t = rX_t dt + \sqrt{\sigma_t}X_t dW_t^0$
- (Volatility Dynamics) $d\sigma_t = \alpha(m - \sigma_t)dt + \beta\sqrt{\sigma_t}dW_t^1$
- (Correlation) $dW_t^1 dW_t^0 = \rho dt$; $|\rho| > 0$, $\rho \in [-1, 1]$

Our option price is now a function of three inputs, denoted $P(t, x, y)$ where y is now in the volatility domain. Remember, our inputs to the price domain and volatility domain are going to be random. As always, we are implicitly assuming P is regular enough to apply Itô's Lemma. We will once again seek to replicate the riskless asset. Take special note that in the original Black-Scholes-Merton model, we required that the stochastic terms of the option differential dP vanish and found an implicit trading strategy in terms of partial derivatives of the option price (Delta Hedging). The Heston Model features two sources of random movement, so we will need a second asset with a different expiry (otherwise it is the same option) in order to properly replicate the riskless asset in this market model. Denote this second asset price $U(t, x, y)$ and assume it is also regular enough to apply Itô's lemma. Denote units of U and X owned over time ϕ_t and Δ_t , respectively. Thus our replicating portfolio value is given by:

$$(4.1) \quad \pi_t = P(t, x, y) + \Delta_t X_t + \phi_t U(t, x, y)$$

We make the same self-financing assumption as before, which is now:

$$(4.2) \quad d\pi_t = dP(t, x, y) + \Delta_t dX_t + \phi_t dU(t, x, y)$$

We have established dX_t in the dynamics assumptions. We can apply the multivariate Itô's lemma to P (we will not restate the inputs here for brevity):

$$(4.3) \quad dP = \partial_t P dt + \partial_x P dX_t + \partial_y P d\sigma_t + \frac{1}{2} \partial_{xx} P dX_t dX_t + \frac{1}{2} \partial_{yy} P d\sigma_t d\sigma_t + \partial_{xy} P dX_t d\sigma_t$$

We have not imposed any additional structure on U yet, so dU will be directly analogous.

Using our box calculus heuristic we can compute the relevant quadratic covariation terms:

- $dX_t dX_t = \sigma_t X_t^2 dt$
- $d\sigma_t d\sigma_t = \beta^2 \sigma_t dt$
- $dX_t d\sigma_t = \beta \rho \sigma_t X_t dt$

Expand (4.2) and rearrange drift terms by the partial derivative they contain, but do not decompose X_t or σ_t into their corresponding drift/diffusion components yet to get:

$$\begin{aligned} d\pi_t &= (\partial_t P + \frac{1}{2} \sigma_t X_t^2 \partial_{xx} P + \beta \rho \sigma_t X_t \partial_{xy} P + \frac{1}{2} \beta^2 \sigma_t \partial_{yy} P) dt \\ &+ \phi_t (\partial_t U + \frac{1}{2} \sigma_t X_t^2 \partial_{xx} U + \beta \rho \sigma_t X_t \partial_{xy} U + \frac{1}{2} \beta^2 \sigma_t \partial_{yy} U) dt \\ &+ (\partial_x P + \phi_t \partial_x U + \Delta_t) dX_t + (\partial_y P + \phi_t \partial_y U) d\sigma_t \end{aligned}$$

Notice the presence of partial derivatives with respect to volatility. The partial derivative of an option with respect to volatility is known as **vega**, one of the famous option Greeks. We can use economic reasoning to deduce that vega is always positive, since a high volatility implies the stock price has fatter tails, making extreme movements into the money more probable. Extreme movements out of the money cannot lose more than what the owner of the option paid for it, but extreme movements into the money can result in theoretically infinite gain, hence the effect of additional volatility on both calls and puts is disproportionately positive.

Now, replicating the riskless asset implies:

$$d\pi_t = r\pi_t dt$$

Both X_t and σ_t are random process components of $d\pi_t$, even if they can be decomposed to get additional drift terms. Thus imposing the replication condition means the randomness coming from dX_t and $d\sigma_t$ must vanish. As in the original Black-Scholes-Merton model, this is going to produce a replicating trading strategy in terms of partial derivatives, but the presence of multiple assets means we will have a system of equations instead of just one. Note that volatility partials will appear in denominators, which are nonzero as we discussed. Our system is given by:

$$(4.4) \quad \phi_t = -\frac{\partial_y P}{\partial_y U}$$

$$(4.5) \quad \Delta_t = -(\phi_t \partial_x U + \partial_x P) = \frac{\partial_y P \partial_x U}{\partial_y U} - \partial_x P$$

For brevity, store the large partial expressions from (4.3) as follows:

$$A = \partial_t P + \frac{1}{2} \sigma_t X_t^2 \partial_{xx} P + \beta \rho \sigma_t X_t \partial_{xy} P + \frac{1}{2} \beta^2 \sigma_t \partial_{yy} P$$

$$B = \partial_t U + \frac{1}{2} \sigma_t X_t^2 \partial_{xx} U + \beta \rho \sigma_t X_t \partial_{xy} U + \frac{1}{2} \beta^2 \sigma_t \partial_{yy} U$$

We then get the compressed expression (after eliminating the stochastic terms):

$$(4.6) \quad d\pi_t = A dt + \phi_t B dt$$

which leads to the replicating condition being:

$$(4.7) \quad A + \phi_t B = r(P + \Delta_t X_t + \phi_t U)$$

Using what we have derived in (4.4) and (4.5), substitute and rearrange to get:

$$(4.8) \quad \frac{A - rP + rX_t \partial_x P}{\partial_y P} = \frac{B - rU + rX_t \partial_x U}{\partial_y U}$$

Notice that the left hand side contains only $P(t, x, \sigma_t)$ terms and the right hand side only contains $U(t, x, \sigma_t)$ terms, along with terms that are in the domain of both. The equality (4.8) thus suggests that both of these expressions are just one function with a time/price/volatility domain. Denote this function $f(t, x, y)$. Following Heston, specify f in the following way:

$$(4.9) \quad f(t, x, y) = -\alpha(m - \sigma_t) + \lambda(t, x, y)$$

Here, λ is called the **Price of Volatility Risk**. Replace the left hand side of (4.8) with $f(t, x, y)$ and expand B fully to get:

$$f(t, x, y) = \frac{\partial_t U + \frac{1}{2} \sigma_t X_t^2 \partial_{xx} U + \beta \rho \sigma_t X_t \partial_{xy} U + \frac{1}{2} \beta^2 \sigma_t \partial_{yy} U - rU + rX_t \partial_x U}{\partial_y U}$$

Then rearrange to get the Heston PDE in terms of the price of the asset U :

$$(4.10) \quad \partial_t U + \frac{1}{2} \sigma_t (X_t)^2 \partial_{xx} U + \beta \rho \sigma_t X_t \partial_{xy} U + \frac{1}{2} \beta^2 \sigma_t \partial_{yy} U \\ - rU + rX_t \partial_x U + [\alpha(m - \sigma_t) - \lambda(t, x, y)] \partial_y U = 0$$

Denote the **log price** $z = \log(x)$ and execute a change of variables from the (t, x, y) domain to the (t, z, y) domain. The Heston PDE for U (which will also describe the behavior of P , just with a different terminal condition) in the transformed domain is now:

$$(4.11) \quad \partial_t U + \frac{1}{2} \sigma_t \partial_{xx} U + (r - \frac{1}{2} \sigma_t) \partial_z U + \beta \rho \sigma_t \partial_{zy} U + \frac{1}{2} \beta^2 \sigma_t \partial_{yy} U \\ - rU + [\alpha(m - \sigma_t) - \lambda(t, z, y)] \partial_y U = 0$$

Following Heston, we will write the market price of risk as a linear function of volatility only: $\lambda(t, z, y) = \lambda y$

4.4. Solving the Heston PDE. With our PDE in hand, we consider the probabilistic interpretation of how options are priced, namely their discounted expected value. Assume for now that the option is a call, denoted:

$$C(t, x, y) = \mathbb{E}[e^{-rT}(X_\tau - \kappa)^+]$$

Execute a change of variables from the (t, x, y) domain to the (T, z, y) domain where $T = (\tau - t)$ denotes time to expiry. Noting that the option is worthless if it expires out of the money, $C(T, Z_t, \sigma_t)$ becomes:

$$(4.12) \quad e^{Z_t} \mathbb{P}^1(T, Z_t, \sigma_t) - e^{-rT} \kappa \mathbb{P}^2(T, Z_t, \sigma_t)$$

Here, $\mathbb{P}^j(T, S_t, \sigma_t)$ each denote the probability of the option expiring in the money at time t conditioned on current price of the underlying, current volatility, and current time to expiry. Putting the general Heston PDE (4.10) in terms of C and executing an easy change of variables transforming the time domain to the time-to-expiry domain, we get the following PDE:

$$(4.13) \quad -\partial_T C + \frac{1}{2} \sigma_t \partial_{zz} C + (r - \frac{1}{2} \sigma_t) \partial_z C + \beta \rho \sigma_t \partial_{zy} C + \frac{1}{2} \beta^2 \sigma_t \partial_{yy} C - rC + [\alpha(m - \sigma_t) - \lambda \sigma_t] \partial_y C = 0$$

We showed in (4.12) that C is a function of the conditional probability functions, so any partial derivative of C can be put in terms of the \mathbb{P}^i terms. We can conveniently summarize the corresponding PDEs by making the appropriate substitutions, setting $\kappa = 0$ with price $X = 1$ to get the PDE for \mathbb{P}^1 and setting $\kappa = -1$ with price $X = 0$ and $r = 0$ to get the PDE for \mathbb{P}^2 . We can then summarize:

$$(4.14) \quad -\partial_T \mathbb{P}^j + \beta \rho \sigma_t \partial_{zy} \mathbb{P}^j + \frac{1}{2} \sigma_t \partial_{yy} \mathbb{P}^j + (r + u_i \sigma_t) \partial_z \mathbb{P}^j + (a - b_i \sigma_t) \partial_y \mathbb{P}^j = 0$$

Here, $u_1 = \frac{1}{2}$, $u_2 = -\frac{1}{2}$, $a = \alpha m$, $b_1 = \alpha + \lambda - \beta \rho$, and $b_2 = \alpha + \lambda$

Now, consider the characteristic function of the random variable $Z_\tau = \log(x_\tau)$. Heston makes the assumption that they are of the following form:

$$f^j(\phi; z, \sigma) = \exp\{C^j(T, \phi) + D^j(T, \phi)\sigma_0 + i\phi z\}$$

By the **Feynman-Kac Theorem**¹⁵, we know that the characteristic functions will also obey the Heston PDE (4.14):

$$(4.15) \quad -\partial_T f^j + \beta \rho \sigma_t \partial_{zy} f^j + \frac{1}{2} \sigma_t \partial_{zz} f^j + \frac{1}{2} \beta^2 \sigma_t \partial_{yy} f^j + (r + u_j \sigma_t) \partial_z f^j + (a - b_i \sigma_t) \partial_y f^j = 0$$

Remember, we are in the time-to-expiry/log-returns/volatility domain. The partials of f^j can be put in terms of C^j and D^j by construction. Putting them in these terms and substituting leads to the following PDEs:

$$(4.16) \quad \partial_T D^j = \beta \rho i \phi D^j - \frac{1}{2} \phi^2 + \frac{1}{2} \beta^2 (D^j)^2 + u_i i \phi - b_i D^j$$

¹⁵A detailed treatment of Feynman-Kac would hinder the accessibility of this paper. However, along with Girsanov's Theorem, it forms the backbone of probabilistic treatments of option pricing

$$(4.17) \quad \partial_T C^j = ri\phi + aD^j$$

To get a solution, we can specify our initial conditions:

$$D^j(0, \phi) = 0$$

$$C^j(0, \phi) = 0$$

With these conditions, (4.16) becomes a **Riccati Equation** and (4.17) becomes an ODE once the Riccati equation is solved. First, the **Heston-Riccati** equation becomes:

$$(4.18) \quad \partial_T D^j = \mathbb{P}^j - Q^j D^j + R(D^j)^2$$

Its steady state solutions are given by the second order ODE:

$$(4.19) \quad \omega'' + Q^j \omega' + P^j R = 0$$

Its general solution is given by:

$$(4.20) \quad D^j = -\frac{1}{R} \frac{\kappa a e^{aT} + b e^{bT}}{\kappa e^{aT} + e^{bT}}$$

Where $a, b \in \mathbb{R}$ are established by initial conditions. Our solution becomes:

$$(4.21) \quad D^j = \frac{(b_i - \beta \rho i \phi + d^j)(1 - e^{d^j T})}{\beta^2 (1 - g^j e^{d^j T})}$$

Where

$$d^j = \sqrt{(\beta \rho i \phi - b^j) - \beta^2 (2u^j i \phi - \phi^2)}$$

and

$$g^j = \frac{b^j - \beta \phi + d^j}{b^j - \beta \rho i \phi - d^j}$$

Plugging this into (4.17) and integrating, we get:

$$(4.22) \quad C^j = \int_0^T ri\phi dy + a \left(\frac{Q^j + d^j}{2R} \right) \int_0^T \frac{1 - e^{d^j y}}{1 - g^j e^{d^j y}} dy + \kappa_1$$

Where κ_1 is a constant. Adding the initial condition $C^i = (0, \phi) = 0$, and substituting d^i, Q^i, g^i , we get the solution for C^i :

$$C^j = ri\phi T + \frac{a}{\beta^2} [(b^j - \beta \rho i \phi + d^j)T - 2 \log\left(\frac{1 - g^j e^{d^j T}}{1 - g^j}\right)]$$

5. APPENDIX: GIRSANOV THEOREM AND FEYNMAN-KAC

While much of the actual pricing machinery we have derived has employed the theory of differential equations, equivalent solutions can often be found using probabilistic reasoning. There are two critical theorems that enable this to happen, which require a bit more exposure to measure theory to fully appreciate. This appendix presents them very briefly to the interested reader. These are adapted from [2].

Theorem 5.1. (*Girsanov's Theorem*)

Let \vec{W}_t be an n -dimensional Brownian Motion on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$.

Let $\vec{\theta}_t$ be another n -dimensional, \mathbb{F} -adapted process. Define a new process:

$$Z_t = \exp\left\{-\sum_{i=0}^n \left[\int_0^t \theta_s^i dW_s^i + \frac{1}{2} \int_0^t (\theta_s^i)^2 ds\right]\right\}$$

Fix a time T and define a single random variable $Z = Z_T$. Then $\mathbb{E}[Z] = 1$ and $Z > 0$ \mathbb{P} -a.s. Further, we can induce a new probability measure \mathbb{Q} on the space by letting Z be its Radon-Nikodym derivative: $\frac{d\mathbb{Q}}{d\mathbb{P}} = Z$. Then the process

$$\widetilde{W}_t^i = W_t^i + \int_0^t \theta_s^i ds$$

is a standard Brownian Motion with respect to \mathbb{Q} for each i .

Theorem 5.2. (*Feynman-Kac Theorem*)

Consider a Stochastic Differential Equation describing a process X_t over a filtered probability space:

$$dX_t = \beta(t, X_t)dt + \gamma(t, X_t)dW_t$$

Let $h(x)$ be some Borel measurable function over our space domain and fix $T > 0$. Define a function over $[0, T] \times X(\Omega)$,

$$g(t, x) = \mathbb{E}^{t,x}[h(X_T)]$$

where expectation is taken conditional on the current time and current state of the process X_t . Then $g(t, x)$ is guaranteed to satisfy the deterministic Partial Differential Equation:

$$\partial_t g(t, x) + \beta(t, X_t)\partial_x g(t, x) + \frac{1}{2}\gamma^2(t, X_t)\partial_{xx}g(t, x) = 0$$

Why are these two theorems important? When we use the abstracted notion of a general probability space, we can possibly make a new probability space with a "fake" probability measure. For instance, the real probability of a fair coin being on heads or tails is 0.5. However, if we define a measure \mathbb{Q} that assigns a measure of 0.75 to heads and 0.25 to tails, \mathbb{Q} is still a valid probability measure, though it is not measuring "real" probabilities of the fair coin. Similarly, stock prices are **not** martingales under the "real" probability measure, but there are other probability measures under which they are. These probability measures are called **Risk Neutral Measures** or **Equivalent Martingale Measures**. In fact, the Fundamental Theorems of Asset Pricing assert that assuming No Arbitrage guarantees the existence of at least one such measure (First Theorem), and assuming completeness of markets guarantees its uniqueness (Second Theorem).

Girsanov's theorem is related to finding alternative probability measures in an abstract probability space and is critical to invoke when working with Equivalent Martingale Measures. As it turns out, we cannot actually get the exact probabilities about where a stock is going to move. Without being able to apply the "real" measure, we have no real hope of computing the "real" expected value of an option on a randomly evolving stock. However, we can compute its expected value under the risk neutral measure, which leads into **Risk Neutral Pricing** arguments. All of our conclusions about Black-Scholes-Merton can be obtained through risk neutral arguments, and Girsanov's theorem is central to making them.

Feynman-Kac relates SDEs to PDEs. The conditional expectation under the risk neutral measure is meant to be an option's price, but it will still be bounded by the terminal condition set at expiry, and will often satisfy some SDE. Feynman Kac enables that risk neutral measure to satisfy a deterministic PDE that may admit solutions through more standard PDE arguments. Together with Girsanov's theorem, this forms the backbone of Risk Neutral Pricing theory.

ACKNOWLEDGMENTS

I would like to thank Beniada Shabani, my advisor, for her extraordinary contribution in helping me build the background to write this paper. Without her efforts I would not have been able to delve nearly as deeply into mathematical finance this summer. I would also like to thank Mark Cerenzia for teaching an incredible class on stochastic processes and stochastic calculus, without which I would not have had the machinery to understand the mathematics needed for this paper in a reasonable amount of time. I also want to thank Peter May for facilitating another incredible year of the REU, and lastly Jitka Stehnova, without whom I could not have stayed in Hyde Park to be a participant in this incredible program.

REFERENCES

- [1] Jean-Pierre Fouque. Multiscale Stochastic Volatility for Equity, Interest-Rate, and Credit Derivatives. 2011.
- [2] Steven E. Shreve. Stochastic Calculus for Finance II: Continuous Time Models. Springer. 2004.
- [3] Mark Cerenzia. Martingales. 2022. Delivered Spring 2022.
- [4] Mike Stecher. Converting the Black-Scholes PDE to The Heat Equation. 2012. URL: <https://www.math.tamu.edu/~stecher/425/Sp12/blackScholesHeatEquation.pdf>
- [5] Fabrice Douglas Rouah. Simplified Derivation of the Heston Model. Fabrice Rouah. 2013.
- [6] Jaska Cvitanic and Fernando Zapatero. Introduction to the Economics and Mathematics of Financial Markets. The MIT Press. 2004.
- [7] Teddy Niemiec. The Black-Scholes-Merton Equations in Practice. University of Chicago REU. 2013.
- [8] <http://assets.press.princeton.edu/chapters/s7834.pdf>
- [9] P. Amster et al. A Black-Scholes option pricing model with transaction costs. Journal of Mathematical Analysis and Applications. 2005.
- [10] Mike Stecher. Options on Dividend Paying Stocks. 2012. URL: <https://www.math.tamu.edu/~mike.stecher/425/Sp12/optionsForDividendStocks.pdf>
- [11] Benoit B. Mandelbrot and Richard L. Hudson. The (mis)Behavior of Markets: A Fractal View of Risk, Ruin, and Reward. Basic Books. 2004.
- [12] Cory Mitchell et al. Volatility Smile Definition and Uses. Investopedia. 2021. URL: <https://www.investopedia.com/terms/v/volatilysmile.asp>
- [13] Lawrence C. Evans. Partial Differential Equations: Second Edition. American Mathematical Society. 2010.
- [14] Kevin Hincks. Trading Volatility: Insights from Professional Traders. TD Ameritrade. 2016. URL: <https://tickertape.tdameritrade.com/trading/mean-reversion-volatility-15463>
- [15] Jinho Bae et al. Why are stock returns and volatility negatively correlated? SSRN Electronic Journal. 2007.
- [16] Lawrence C. Evans. Partial Differential Equations: Second Edition. American Mathematical Society. 2010.