2012-06-18

Spread Option Pricing with Stochastic Interest Rate

Yi Luo
Brigham Young University - Provo

Spread Option Pricing with Stochastic Interest Rate

Yi Luo

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Kening Lu, Chair
John Dallon
Christopher Grant
Jeff Humpherys
Kenneth Kuttler

Department of Mathematics

Brigham Young University

August 2012

ABSTRACT

Spread Option Pricing with Stochastic Interest Rate

Yi Luo
Department of Mathematics, BYU
Doctor of Philosophy

In this dissertation, we investigate the spread option pricing problem with stochastic interest rate. First, we will review the basic concept and theories of stochastic calculus, give an introduction of spread options and provide some examples of spread options in different markets. We will also review the market efficiency theory, arbitrage and assumptions that are commonly used in mathematical finance. In Chapter 3, we will review existing spread pricing models and term-structure models such as Vasicek Mode, and the Heath-Jarrow-Morton framework. In Chapter 4, we will use the martingale approach to derive a partial differential equation for the price of the spread option with stochastic interest rate. In Chapter 5, we will study the spread option numerically. We will conclude this dissertation with ideas for future research.

Keywords: Spread Option, Stochastic Interest Rate, Vasicek Model, Term Structure, HJM Frame Work, GMM

Acknowledgments

I would like to express my deep appreciation to my advisor and mentor Kening Lu for the time, guidance, encouragement he has given me, for his patience and for the opportunities he has provided for me to obtain a PhD degree at BYU.

I would like to thank my grandparents Zhensheng Luo, and Yunxiang Li for teaching me the value of work, and my parents Yongjiang Luo and Meiju He, for their encouragement, concern and support throughout college and graduate school, I also thank my aunt Jianping Luo and uncle Yafei He for their help to make my learning experience at BYU possible. Lastly, I also thank my wife Sarah, for her support and encouragement for my research and patience while I was busy working.

I would also like to extend my appreciation to my committee: Professor John Dallon, Christopher Grant, Jeff Humpherys, and Kenneth Kuttler for their willingness to give of their time and provide counsel. I thank the graduate committee for admitting me and the financial support they provided.

Contents

List of Tables

List of Figures

Chapter 1. Introduction

Every transaction in finance can be viewed as buying or selling a risk. For example, when the bank sells a mortgage to another bank, it is selling the risk that the loan could default, and the counter party of this transaction is buying the risk of default. The success of an investment strategy is determined by the amount of the return given a certain level of risk. A high amount of risk is generally associated with the possibility of high return, and vice versa. This relationship is one of the most important aspects in forming an investment strategy, so we need to find a way to effectively manage risk when investing. In order to accomplish this goal, we first need to achieve an understanding of what risk is in finance.

In finance, there are two types of assets: risk-free assets and risky assets. The risk-free asset can be defined as the asset with determined future value, such as a government bond. When we buy a government bond, we have a guarantee that we will receive a series of coupon payments at fixed future dates until the bond matures. Another example of a risk-free asset is a savings account since you are guaranteed to get back the initial deposit plus any interest.

A risky asset is an asset that is not risk-free. It has uncertain future value. The source of the risk comes from this uncertainty. An example of a risky asset is a share of the stock of a publicly owned company because the movement of the stock price is undetermined. Another example is when we buy or sell a foreign currency since in this case we expose ourselves to the risk of movement on the exchange rate.

As mentioned above, the prevalence of risk in finance is so substantial that we need to find a way of effectively managing it. Risk management includes identifying and assessing risk, controlling the impact of unfortunate events, and maximizing the realization opportunities. The purpose of financial mathematics is to identify risk by studying the relationships between various assets in the financial market. As mentioned before, risks arise from uncertainty, or more precisely, risks arise from the uncertainty of the dependence structure between market variables. For instance, stocks can go up and down together since the rise of one stock might lead to the rise of similar stocks. Markets can go up and down together as well: a crash in the US housing market in 2007 lead to the crash of the US stock market, which then lead to the ongoing global financial crisis. In contrast, markets can also move in opposing directions. Since the market crash at the end of 2007, the gold market has approximately doubled, with the price of gold going from 833 dollars per ounce to 1655 per ounce. This shows that if we can understand how the market variables are correlated, then we can manage risk easily. Unfortunately, market variables are typically not correlated in a deterministic way. This is why stochastic methods are a natural choice for modeling the correlation between market variables.

Two key issues that arise when working to manage risk in finance are how to quantify the correlation of market variables and how to hedge the risk arising from this correlation. The connection of these two important issues is the class of derivatives called *spread options*. In general, an option is a contract that gives its owner the right to buy or sell an asset at a fixed price (called the *strike price*) in the future. A *call option* gives the owner the right to buy, and a *put option* gives the owner the right to sell. A spread option is an option written on the spread in the price of two underlying assets S_1 and S_2. The payoff for a call spread option with strike price K is given as

$$(S_{1,T} - S_{2,T} - K)^+ = \max(S_{1,T} - S_{2,T} - K, 0) \tag{1.1}$$

The two prices processes $S_{1,t}$ and $S_{2,t}$ forming the spread can be referred to as asset prices, future prices, indexes, yields, or exchange rates. Spread options are mostly traded on the Over-the-counter market. Exchanges such as the New York Mercantile Exchange (NYMEX) and Chicago Board of Trade (CBOT) trade on energy and commodity spread options. Other types of spread options in various markets include the Note Over Bond spread, Treasury Bills-Eurodollar spread, and Municipal Over Bonds spread in the fixed income market, calendar spread and crush spread in the commodity market, and spark spread in the energy markets. Compared to options on a single asset, spread options not only can hedge movement on one asset, but they also capture the correlations between two assets, or indexes. This allows investors to hedge or speculate more risk. Despite the apparent usefulness of spread options in hedging and speculating, the pricing of such an option is quite difficult. The difficulty mainly comes from two parts, the first part is what mathematical model to use to describe the prices of the underlying processes $S_{1,t}$ and $S_{2,t}$, and the second part is what discount factor we should use to discount the future payoff.

In his PhD thesis in 1900, Louis Bachelier created a model using Brownian motion to show the fluctuations in stock prices. He argued that the small change in price over a short time should be independent of the current value of the price. He also proposed a model for the price of a stock now known as the arithmetic Brownian motion. His thesis provided a foundation for many subsequent developments in probability theory and stochastic analysis for the next 65 years. Several mathematicians including Wiener, Kolmogorov, Doob, and Ito carried out these developments. In 1965, Paul Samuelson introduced the geometric Brownian motion model for the price of an asset. In 1973, Black, Scholes, and Merton adopted Samuelson's price model and developed the Black-Scholes formula for the price of the European call option. They also introduced the idea of arbitrage free and risk neutral pricing.

Following the work of Black, Scholes and Merton, the insight to adapt Bachelier's arithmetic Brownian motion was first exploited in a paper by Wilcox. In 1990, Wilcox [1] derived a closed form formula for spread options by adapting the arithmetic Brownian motion to the price process for the underlying assets. However, as mentioned in [2], the formula is not arbitrage-free, and Poitras proposed a two factor model by modeling the two underlying assets separately with arithmetic Brownian motion and then derived a closed form formula for the spread option. In 1994, Shimko used a double-integral approach involving the geometric Brownian motion for the asset prices. Even though he did not come up with a closed form formula, Shimko's work inspired other researchers to find a way to get a closed form solution. Under the geometric Brownian motion model, only an exchange option, which is a spread option with strike price zero, has a closed form solution by using one asset as a numeraire.

In [3], Margrabe derived a closed form formula with a PDE approach.

In most of the previous work on spread option pricing, the interest rate was assumed to be constant, which is far from reality. There are many stochastic models for interest rates, such as Vasicek model, Cox, Ingersoll and Ross model, and Hull-White model. In [4], Liu and Wang derived a closed form formula for the exchange option when interest rate is stochastic. In this dissertation, we first derive a partial differential equation for the price of a call spread option with the stochastic interest rate. We will also study the price from a numerical standpoint, specifically how the price changes along with different values of the parameters in the interest rate model. From the numerical results, we will see that under the stochastic interest rate model, the price decreases as the mean reverting speed increases, and it increases when the long term mean increases. We will conclude the dissertation with ideas for future research.

Chapter 2. Preparation

In this chapter, we will review basic concept and theories of stochastic calculus. We will also provide some examples of spread options that are traded in different markets. In the end of the chapter, we will review the effcient market hypothesis, definition of arbitrage and assumptions in mathematical finance.

2.1 Stochastic Calculus

In this chapter, we will review basic concepts and results on the Brownian motion and Ito stochastic calculus. We refer readers to [5] and [6] for the details of the proofs and some of definitions.

2.1.1 The Brownian Motion. First we will define what is a stochastic process:

Definition 2.1. A stochastic process is a parameterized collection of random variables

$$\{X_t\}_{t\in[0,T]}$$

defined on a probability space $(\Omega, \mathcal{F}, P)$.

Note that for each fixed $t \in [0,T]$ we have a random variable $X_t(\omega)$, and for a fixed $\omega \in \Omega$ we have a function that maps t to $X_t(\omega)$ which is called a path of X_t. We can also regard the process as a function of two variables $(t,\omega) \to X(t,\omega)$. A good example of a stochastic process is the Brownian motion.

Definition 2.2. A real-valued stochastic process $W(t), t \in [0,T]$ is called a Brownian motion if

- the process has independent increments for $0 \le t_0 \le t_1 \le \cdots \le t_n \le T$
- for all $t \ge 0, h > 0$, $W(t+h) - W(t)$ is normally distributed with mean 0 and variance h,

- the function $t \to W(t)$ is continuous a.s.

We also call a Brownian motion a Wiener processes. Here are some properties of a Brownian motion

(i) W_t is continuous in t a.s.

(ii) For any interval $[a, b] \subset [0, \infty)$, W_t is not monotone.

(iii) W_t is not differentiable at any point.

2.1.2 Filtration and Martingale Property. Information is one of the most important piece in finance. So in mathematic finance, we need some notation for the amount of available at a time, and we call it a filtration. In mathematical finance, when we use Brownian motion to describe the "randomness" of the market or an asset, the information available at time t is totally determined by the Brownian motion $W(t)$. Thus we use the filtration for Brownian motion to describe the amount of information.

Definition 2.3. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and $W(t)$ be a Brownian motion on this space. A filtration for the Brownian motion is a collection of $\sigma-$algebra $\{\mathcal{F}_t\}_{t=0}^{\infty}$, such that

(i) $\mathcal{F}_t \subset \mathcal{F}$ for all t.

(ii) For $0 < s < t$, we have $\mathcal{F}_s \subset \mathcal{F}_t$.

(iii) The Brownian motion W_t is adapted to $\mathcal{F}_t$ for all $t \geq 0$, i.e. W_t is $\mathcal{F}_t$ measurable.

(iv) For $0 < s < t$, the increment $W_t - W_s$ is independent of $\mathcal{F}_s$.

Properties (ii) and (iii) guarantee that we can get the information at time t by observe the movement of Brownian motion up to time t. Property (iv) implies that we can not use the today's information to predict the feature, which leads to the efficient market hypothesis.

One of the assumption in financial mathematics is there is no arbitrage in the market. This implies that investing is a "fair game", which means that on average, we are not expected to have gains or losses. If we translate this into mathematics, we say it is a martingale or has martingale property. Recall the definition of a martingale:

Definition 2.4. Let $(\Omega, \mathcal{F}, \mathbb{P})$ to a probability space, and $\{\mathcal{F}_t\}$ is a filtration for this space ,and X_t is a stochastic process on this space. X_t is call a martingale process, if

(i) X_t is adapted to $\mathcal{F}_t$ for all $t \geq 0$

(ii) $E[|X_t|] < \infty$ for all $t \geq 0$

(iii) For $0 < s < t$, we have $E[X_t|\mathcal{F}_s] = X_s$

If we consider X_t as an investment portfolio, then the third property says that with the information available now (at time s), on average the investor is not going to make any profit or have any losses in a future time t. Thus if the price process for every asset in the market is martingale, then there is no arbitrage. To see why, let's start with a portfolio X_t with zero initial value, $X_0 = 0$. Since a portfolio is a linear combination of assets, then by the linearly of conditional expectation, X_t is also a martingale. Thus for any $t > 0$, we have

$$E[X_t|\mathcal{F}_0] = E[X_t] = X_0 = 0 \tag{2.1}$$

and recall that an arbitrage portfolio means that for any $t > 0$, we have

$$\mathbb{P}(X_t \geq 0) = 1 \quad \text{and} \quad \mathbb{P}(X_t > 0) > 0 \tag{2.2}$$

clearly this will never happen if X_t is martingale.

Theorem 2.5. *Brownian motion W_t is a martingale.*

2.1.3 Ito Integral.

Suppose we have a model with some random noise

$$\begin{cases} \dfrac{dX_t}{dt} = a(t, X_t) + b(t, X_t) \cdot \xi \\ X(0) = x_0 \end{cases} \tag{2.3}$$

Ito considered the case where the noise term $\xi = \Delta W_t$, and W_t is Brownian motion. Then we have the stochastic differential equation

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t, \tag{2.4}$$

a stochastic process X_t is a solution of (2.4) if

$$X_t = x_0 + \int_0^t a(s, X_s)ds + \int_0^t b(s, X_s)dW_s \tag{2.5}$$

Note that a Brownian motion is nowhere differentiable, so we need to define the term

$$\int_0^t b(s, X_s)dW_s.$$

Definition 2.6. Suppose $0 \le S \le T$, let $\mathcal{D} = \mathcal{D}(S, T)$ be the class of functions that

$$f(t, \omega) : \ [0, \infty] \times \Omega \to \mathbb{R}^n, \tag{2.6}$$

satisfy

(i) The function $(t, \omega) \to f(t, \omega)$ is $\mathcal{B} \times \mathcal{F}$ measurable, where $\mathcal{B}$ is the Borel algebra.

(ii) f is adapted to $\mathcal{F}_t$.

(iii) $E\left[\int_S^T f(t, \omega)^2 dt\right] < \infty$.

A function $\phi \in \mathcal{D}$ is called an elementary function if

$$\phi(t,\omega) = \sum_{j=1}^{k-1} e_j(\omega)\chi_{[t_j,t_{j-1})}(t) \tag{2.7}$$

where $e_j(\omega)$ is $\mathcal{F}_{t_j}$ measurable, and $\chi_{[t_j,t_{j-1})}(t)$ is the indicator function, (in the measure theory χ is called the characteristic function). Then the **Ito integral of elementary functions** is defined as following:

Definition 2.7. Let $\{S = t_1 < t_1 < \cdots < t_k = T\}$ is a partition of the interval $[S,T]$, then

$$\int_S^T \phi(s,\omega)dW_t(\omega) = \sum_{j=1}^{k-1} e_j(\omega)(W_{s_{j+1}}(\omega) - W_{t_j}(\omega)) \tag{2.8}$$

Theorem 2.8. ***Properties of Ito Integral for elementary function*** *The Ito integral for elementary function defined as in* (2.7)*, has following properties:*

(i) For each t, the Ito integral is $\mathcal{F}_t$-measurable

(ii) Let $\phi(t)$ and $\psi(t)$ be two bounded elementary function, and α be a constant, then we have

$$\alpha\int_S^T \phi(t)dW_t \pm \int_S^T \psi(t)dW_t = \int_S^T \alpha\phi(t) \pm \psi(t)dW_1 \tag{2.9}$$

(iii) Let $X_t = \int_S^t \phi(u)dW_u$, then, X_t integral is martingale for all $t \in [S,T]$.

(iv) $E\left[\int_S^t \phi(u)dW_u\right] = 0$ for all $t \in [S,T]$.

Now we introduce a important result of Ito integral for elementary functions:

Lemma 2.9. ***(The Ito isometry of Ito integral for elementary functions)*** *If $\phi(t,\omega)$ is bounded and elementary then*

$$E\left[\left(\int_S^T \phi(t,\omega)dW_t(\omega)\right)^2\right] = E\left[\int_S^T \phi^2(t,\omega)dt\right] \tag{2.10}$$

Now we can extend this result from elementary functions to all functions in $\mathcal{D}$. It can be done in the following three steps, for more details please read [5]:

Step 1. Let $g \in \mathcal{D}$ be bounded and continuous in t. Define

$$\phi_n = \sum_j g(t_j, \omega)\chi_{[t_j,t_{j+1})}(t) \tag{2.11}$$

Then ϕ_n are elementary functions since $g \in \mathcal{D}$, and is also bounded. Because g is continuous in t, then it is uniformly continuous in $[S, T]$, thus for a given ϵ, there exists δ such that when $|t_j - t_i| \leq \delta$ we have

$$|g(t_i, \omega) - g(t_j, \omega)| \leq \epsilon. \tag{2.12}$$

Now we make the interval $[t_j, t_{j+1})$ have length less than δ. Then for all $t \in [t_j, t_{j+1})$ we have

$$\int_S^T (g - \phi_n)^2 dt = \sum_k \int_{t_j}^{t_{j+1}} (g - \phi_n)^2 dt = \epsilon^2 (T - S). \tag{2.13}$$

This implies that for each ω we have $\int_S^T (g - \phi_n)^2 dt \to 0$ as $n \to \infty$, thus by the bounded convergence theorem we have

$$E\left[\int_S^T (g - \phi_n)^2 dt\right] \to 0, \quad \text{as } n \to \infty$$

Step 2. Let $h \in \mathcal{D}$ be bounded, $|h| \leq M$. Let ψ_n be a continuous function in t, such that

(1) $\psi_n(t) \geq 0$

(2) $\psi_n(t) = 0 \quad$ for $t \leq -\frac{1}{n}$, or $t \geq 0$

(3) $\int_{-\infty}^{\infty} \psi_n(t) dt = 1$.

Now define

$$g_n(t, \omega) = \int_0^t \psi_n(s - t) h(s, \omega) ds \tag{2.14}$$

then g_n is continuous in t and bounded. With the same reasoning as in step 1 we have $\int_S^T (h - g_n)^2 dt \to 0$ as $n \to \infty$, thus by the bounded convergence theorem we have

$$E\left[\int_S^T (h - g_n)^2 dt\right] \to 0, \quad \text{as } n \to \infty$$

Step 3. Let $f \in \mathcal{D}$, then define

$$h_n(t,\omega) = \begin{cases} -n & f(t,\omega) < -n \\ f(t,\omega) & -n \leq f(t,\omega) \leq n \\ n & f(t,\omega) > n \end{cases} \tag{2.15}$$

Then $h_n \in \mathcal{V}$, and is bounded by $|f|$, then by the Dominated Convergence Theorem we have

$$E\left[\int_S^T (f - h_n)^2 dt\right] \to 0 \quad \text{as } n \to \infty \tag{2.16}$$

Now we are ready for the definition of the Ito integral.

Definition 2.10. (**The Ito integral**) Let $f \in \mathcal{D}(S,T)$, Then the Ito integral of f is defined by

$$\int_S^T f(t,\omega) dW_t(\omega) = \lim_{n\to\infty} \int_S^T \phi_n(t,\omega) dW_t(\omega) \tag{2.17}$$

where $\{\phi_n\}$ is a sequence of elementary functions such that

$$E\left[\int_S^T (f - \phi)^2 dt\right] \to 0 \quad \text{as. } n \to \infty. \tag{2.18}$$

We can choose such a sequence by steps 1-3. Hence, we should have following properties for the Ito integral with general integrands [5], [6]

Theorem 2.11. ***Property for Ito integral with general integrands*** *Let $f, g \in \mathcal{D}(0,T)$, and $0 \leq s < t < r < T$, a is a constant, then*

(i) $\int_s^r f(u) dW_u = \int_s^t f(u) dW_u + \int_t^r f(u) dW_u$

(ii) $\int_s^t af(u) \pm g(u)dW_t = a\int_s^t f(u)dW_u \pm \int_s^t g(u)dW_u$

(iii) $\int_s^t f(u)dW_u$ *is* $\mathcal{F}_t$*-measurable*

(iv) $E\left[\int_s^t f(u)dW_u\right] = 0$

(v) Let $X_t = \int_0^t f(u)dW_u$*, then* X_t *is martingale.*

Proof. These properties are true for elementary functions, thus we can use steps 1-3 to obtain them for $f, g \in \mathcal{D}$. □

We can also have the **Ito isometry** for all functions in $\mathcal{D}$ as:

Theorem 2.12. ***(The Ito isometry)*** *For each* $f \in \mathcal{D}$ *we have*

$$E\left[\left(\int_S^T f(t,\omega)dW_t(\omega)\right)^2\right] = E\left[\int_S^T f^2(t,\omega)dt\right] \tag{2.19}$$

2.1.4 Ito Formula.

Theorem 2.13. ***(The 1-dimensional Ito formula)*** *Let* X_t *be a stochastic process given by*

$$dX_t = u(t, X_t)dt + v(t, X_t)dW_t \tag{2.20}$$

Let $g(t,x) \in C^2([0,\infty] \times \mathbb{R})$*, then*

$$Y_t = g(t, X_t) \tag{2.21}$$

is a stochastic process and

$$dY_t = \frac{\partial g}{\partial t}(t, X_t)dt + \frac{\partial g}{\partial x}(t, X_t)dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t, X_t) \cdot (dX_t)^2 \tag{2.22}$$

where $(dX_t)^2 = (dX_t) \cdot (dX_t)$ *is computed according to the rules*

$$dt \cdot dt = dt \cdot dW_t = dW_t \cdot dt = 0, \quad dW_t \cdot dW_t = dt \tag{2.23}$$

Proof. Please refer to [5] for a sketch of the proof. □

Note that since W_t is nowhere differentiable, dW_t is a just a notation, but with the meaning of (2.23). I will provide an explanation from [6], first let us look at the quadratic variation of Brownian motion W_t. Recall the definition of quadratic variation:

Definition 2.14. Let $f(t)$ be a function defined on $[0, T]$, the quadratic variation of f up to time T is

$$[f, f](T) = \lim_{\|\Pi\| \to 0} \sum_{i=0}^{n-1} (f(t_{i+1}) - f(t_i))^2 \tag{2.24}$$

where $\Pi = \{0 = t_0 < t_1 < \cdots < t_n = T\}$ and $\|\Pi\| = max_{i=0,1,\ldots,n-1}\{t_{i+1} - t_i\}$

We have following theorem in [6] regarding the quadratic variation of Brownian motion

Theorem 2.15. *Let W_t be a Brownian motion, then $[W, W](T) = T$ for all $T \geq 0$ a.s.*

Proof. Let $\Pi = \{0 = t_0 < t_1 < \cdots < t_n = T\}$, and $\|\Pi\| = max_{i=0,1,\ldots,n-1}\{t_{i+1} - t_i\}$, then let

$$Q_\Pi := \sum_{i=0}^{n-1} (W_{i+1} - W_i)^2 \tag{2.25}$$

where $W_i := W_{t_i}$. Note that Q_Π is a random variable. We have

$$E[Q_\Pi] = \sum_{i=0}^{n-1} E[(W_{i+1} - W_i)^2] = \sum_{i=0}^{n-1} (t_{i+1} - t_i) = T \tag{2.26}$$

but this is not enough to say that $Q_\Pi \to T$ as $\|\Pi\| \to 0$ a.s. We have to look at the variance of Q_Π.

$$\begin{aligned}
Var(Q_\Pi) &= \sum_{i=0}^{n-1} Var((W_{i+1} - W_i)^2) \\
&= \sum_{i=0}^{n-1} \left(E[(W_{i+1} - W_i)^4] - (t_{i+1} - t_i)^2\right) \\
&= \sum_{i=0}^{n-1} 2(t_{i+1} - t_i)^2 \\
&\leq \sum_{i=0}^{n-1} 2\|\Pi\|(t_{i+1} - t_i) \\
&= 2\|\Pi\|T
\end{aligned} \tag{2.27}$$

and it goes to zero as we refine the partition. Thus we can say that Q_Π goes to T a.s. $\square$

As remarked in [6], for $0 < T_1 < T_2$, we have $[W,W](T_1) = T_1$ and $[W,W](T_2) = T_2$, thus we have $[W,W](T_2) - [W,W](T_1) = T_2 - T_1$. This implies that Brownian motion accumulates one quadratic variation per unite time and we write $dW_t dW_t = dt$ to record this fact.

Similarly, we can show that for a partition Π,

$$\lim_{\|\Pi\| \to 0} \sum_{i=0}^{n-1} (W_{i+1} - W_i)(t_{i+1} - t_i) = 0 \tag{2.28}$$

and

$$\lim_{\|\Pi\| \to 0} \sum_{i=0}^{n-1} (t_{i+1} - t_i)^2 = 0 \tag{2.29}$$

which means that the function $f(t) = t$ accumulates no quadratic variation per unit time and two functions t and W_t together also accumulate no quadratic variation per unit time, thus we can record these two facts by $dW_t dt = dtdt = 0$.

The Ito formula also gives:

Theorem 2.16. ***(Integration by parts)*** *Suppose $f(t,\omega)$ is continuous and of bounded* *variation with respect to $s \in [0,t]$ for a.a.ω. Then*

$$\int_0^t f(s)dW_s = f(t)W_t - \int_0^t W_s df_s. \tag{2.30}$$

Proof. Let $Y_t = f(t)W_t$, then by Ito formula, we have

$$d(f(t)W_t) = W_t f'(t)dt + f(t)dW_t \tag{2.31}$$

integrate this equation from 0 to t, we will get the desired result. $\square$

The Ito formula has a multi-dimensional version:

Theorem 2.17. ***(Multi-dimensional Ito formula)*** *Let*

$$dX_t = u(t, X_t)dt + v(t, X_t)dW_t \tag{2.32}$$

be an n-dimensional Ito process as above. Let $g(t,x) = (g_1, \cdots, g_n)$ be a C^2 map from $[0,\infty) \times \mathbb{R}^n$ in to $\mathbb{R}^m$, Then $Y(t,\omega) = g(t, X_t)$ is an stochastic process and for $k = 1, \cdots, n$

$$dY_k = \frac{\partial g_k}{\partial t}(t, X_t)dt + \frac{\partial g_k}{\partial x_i}(t, X_t)dX_i + \frac{1}{2}\sum_{i,j}\frac{\partial^2 g_k}{\partial x_j \partial x_j}(t, X_t) \cdot dX_i dX_j \tag{2.33}$$

We also have the Ito product formula

Theorem 2.18. ***(Ito product formula)*** *If X_t, Y_t are Ito processes, then*

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + dX_t dY_t \tag{2.34}$$

Proof. Let $g(x,y) = x \cdot y$, then $X_t Y_t = g(X_t, Y_t)$, by the Ito formula we have

$$\begin{aligned} d(X_t Y_t) &= \frac{\partial g}{\partial X_t}dX_t + \frac{\partial g}{\partial Y_t}dY_t + \frac{1}{2}\frac{\partial^2 g}{\partial X_t^2}(dX_t)^2 + \frac{1}{2}\frac{\partial^2 g}{\partial Y_t^2}(dY_t)^2 + \frac{\partial^2 g}{\partial X_t \partial Y_t}dX_t dY_t \\ &= Y_t dX_t + X_t dY_t + 0 + 0 + dX_t dY_t. \end{aligned} \tag{2.35}$$

Therefore

$$X_t dY_t = X_t Y_t + Y_t dX_t + dX_t dY_t$$

$\square$

2.2 Spread Option

The word spread is often understood as the difference between the bid and ask prices, but it is also used for the difference between two indexes, for example: the spread of the prices of two stocks, the spread of the yields of two Treasury bonds with different maturity, the spread between prices of feature contracts with different deliver dates. A spread option by definition is an option written on the difference between the values of two assets. The choices of two assets is limitless, and can across all types of markets. Most of spread option trading are done in the over the counter market, and only small amount are traded in an exchange, energy spreads are traded in The New York Mercantile Exchange (NYMEX) for example and some commodity spreads are traded in the Chicago Board of Trade (CBOT).

2.2.1 Spread Option Trading in Different Markets. Spread options are traded mostly in the fixed income markets, and commodity market. In this section, we will give a brief review on different types of spread options that are traded in each market.

Fixed Income Markets. In the US fixed income market, most traded spread options are yield spread options. It is an option written on the difference between the yields-to-maturity of two bonds or so called two debt obligations. The two bonds can have different maturities, credit risk, etc.

The *Note Over Bond* spread (NOB for short) is a spread option created by taking opposite positions in feature contracts of 30-year treasury bonds with positions in feature contracts of 10-year treasury notes. It is studied in [7]. The *Treasury Bills - Eurodollar* spread option is written on the difference between the three-month LIBOR and the three-month T-bill interest rate. The size of the spread is usually in basis points. The 3-month T-bill rate is usually considered as a risk-free rate and the LIBOR reflects the credit standing of commercial banks. An increasing on the spread indicates that the investors are risk-averse. This is an example of credit spread. The *Municipal Over Bonds* spread (MOB) is the difference between the yields of a municipal bond and a treasury bond with the same time-to-maturity.

Commodity Market. In the commodity market, spread options are usually created by taking long or short positions in forward or feature contracts. They are usually referenced as futures spread.

A *calendar spread* refers to an option taking position in two contracts with different expiration date for the same commodity. An example of calender spread is given in [8], a spread between the spot and three-month prices of copper on the London Metal Exchange (LME).

Another example of spread options in Commodity Market is *crush spread option*. It is traded on the CBOT. The underlying assets are soybean, soybean oil and soybean meal. It is created by taking a long position of futures on soybean and taking a short position on futures of soybean oil and soybean meal. This option indicates the difference between the cost of a commodity and the combined sales revenue of its finished derivative products, this is called the gross processing margin. A soybean meal or oil producer can use this option to hedge the movement on the price of his product and the price of soybeans. This kind of spread is called product spread.

Oil, gas and electricity are also traded in a commodity market, more specifically, in the energy market. One of the most traded spread options in energy markets is *crack spread*. It is created by simultaneously taking a long position in future contracts in crude oil and taking short position in its refined petroleum products such as gasoline, heating oil, diesel, etc. The most popular crack option is called the $3:2:1$ crack option. The option consists of three crude oil futures, two gasoline futures, and one heating oil future [9].

Similar to the crack spread option, the spark spared option is also another example in the energy market. The spark spread involves the simultaneous purchase and sale of electricity and natural gas futures contracts. The spread is expressed in dollar per mega Watt hour. Since the price for natural gas is expressed in dollar per British thermal unit (btu), we need to use the heat rate to convert the units. The heat rate measures how may btu of natural gas is need to generate one Kilo Watt hour electricity. Let p_{ng} denote the prices of natural gas, p_e denote the price of electricity and r is the heat rate, then the spread is calculated as:

$$s = p_{ng} \times r \times 1000 - p_e \tag{2.36}$$

In [10], a closed formula of the price of such spread option is given under both geometric brownian motion model and mean-reversion model with strike price is zero.

2.3 Market Efficiency, Arbitrage and Assumptions

In this section, we will review efficient market hypothesis, definition of arbitrage and assumptions in mathematical finance.

2.3.1 Efficient Market Hypothesis. Efficient Market Hypothesis (EMH) gives an assumption on the relation between an asset's price and the information about that asset. The general statement of the EMH is: all available information of an asset is already included in the asset's price. There are three version of the EMH, the difference among these versions is what is meant by "all available information".

The **Strong-form** of EMH states that all relevant information including inside information is included in the asset's prices. This assumption is quite extreme and not realistic since inside trading is prohibited by the SEC.

The **Semistrong-form** of EMH assets that the asset's price includes all the public data of the asset. That includes historical data, if the asset is a stock, the price includes the earning forecast, quality of management, balance sheet information, etc. The bottom line is that if an investor can get a piece of information from public resources, then it is already included in the price.

The **Weak-form** of EMH states that all the historical data is included in its price. This form implies that studying past prices data for an asset and trying to find some kind of trend is pointless. This form of efficiencies is also refered to as Markov property and financial mathematics is based on this form of EMH.

2.3.2 Arbitrage. Another fundamental concept in financial mathematics is arbitrage. The mathematical definition of arbitrage is:

Definition 2.19. Let $X(t)$ be a stochastic process, which represents the value of a portfolio with $X(0) = 0$. An arbitrage is that for some time $T > 0$, we have

$$\mathbb{P}(X(T) \geq 0) = 1 \quad \text{and} \quad \mathbb{P}(X(T) > 0) > 0 \tag{2.37}$$

In an arbitrage free market,if two portfolio A and B has same value in any time in the feature, then they should have same value now. In the next section, we should see how to use this property of arbitrage free to price a derivative.In financial mathematics, we look for the price in an arbitrage-free market or the price that will not produce an arbitrage opportunity, and we call this the arbitrage-free price.

2.3.3 Assumptions in financial mathematics. In financial mathematics, we generally assume following:

(i) Not moving the market. We usually assume that the trade we made will not move the market prices of the assets.

(ii) No transaction cost. When we price a derivative, we assume we can trade with no transaction cost.

(iii) Fraction shares. In financial mathematics, we usually don't restrict the number of share of stock or derivative to be integers, but all real numbers. This means that we are assuming we can sell or buy half share of a stock. One certainly cannot do this, but if we trade in terms of millions, then this makes sense.

(iv) Short sell. In the process of developing a hedging strategy, the number of assets in a portfolio is assumed to be all real numbers, which means we can short sell our assets at will.

Chapter 3. Mathematical Modeling

This chapter reviews the existing models to give us a general frame work on the approaches to the spread option pricing problem. The framework can be broke down into following steps:

- identify the stochastic model for the price of the two underlying assets
- write down the stochastic model in the risk-neutral measure
- Calculate the expectation of discounted payoff.

In all the models we shall see below, the interest rate is assumed to be constant. This assumption is far away from reality. The new model we propose in this dissertation allows the interest rate to be stochastic. In the section 2, we will review the term-structure models that are used for the stochastic interest rate for this dissertation. In section 3, we will review the the relation between the short rate model and the HJM framework.

3.1 Existing Models for Pricing Spread Option

In this section, I will review existing models for pricing spread option. Here we consider a spread option with two underlying assets, their prices are denoted as $S_1(t), S_2(t)$, which are two stochastic process adapted to the filtration $\mathcal{F}_t$. We shall focus on pricing a European call on the spread $S_1(t) - S_2(t)$ with a strike price K, and maturity T. The payoff of such an option at maturity T is $(S_1(T) - S_2(T) - K)^+$.

In order to price a spread option, the first step is to determine the number of independent sources of uncertainty or randomness thatare in the model. For spread option, it is the spread on the prices of the two underlying assets. We can either only consider the spread $S_1(t)-S_2(t)$ as a signal process, such models are called one-factor model, or consider the prices of two underlying asset separately, which are called two-factor model.

3.1.1 One Factor Models.

In a one factor model, we consider the spread of two underlying prices as a single stochastic process or stochastic factor. This approach, is relatively simple compare to the two factor models, and only requires one set of historical data for price and volatility. In addition, one factor models do not require any knowledge of the correlation of two underlying assets.

Geometric Browinian Motion. The simplest case of one factor model is assume that the spread $S_t = S_1(t) - S_2(t)$ satisfies the Geometric Brownian Motion:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{3.1}$$

where dW_t is the standard Brownian motion. Then we can solve for S_t get

$$S_t = S_0 \exp\left\{\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W_t\right\} \tag{3.2}$$

In this case, the call spread option can be priced as a European call option using Black-Scholes formula

$$V(t) = N(d_1)(S_1(t) - S_2(t)) - N(d_2) k e^{-r(T-t)} \tag{3.3}$$

where

$$\begin{aligned} d_1 &= \frac{\ln \frac{S_1(t)-S_2(t)}{K} + \left(\mu_s + \frac{\sigma_s^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}} \\ d_2 &= \frac{\ln \frac{S_1(t)-S_2(t)}{K} + \left(\mu_s - \frac{\sigma_s^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}} \end{aligned} \tag{3.4}$$

If both underlying assets are tradeable, such as stocks, then we can think of the spread also as a tradeable asset, thus the parameter $\mu_s = r$.

In order to calculate the parameter σ_s. we can use the historical data for each underlying asset to get the variance, then using following equation

$$\sigma_s = \sqrt{\sigma_{S_1}^2 + 2\rho\sigma_{S_1}\sigma_{S_2} + \sigma_{S_2}^2} \tag{3.5}$$

This implies that the difference of two correlated lognormal random variables is still lognormal, which is false as pointed in [11].

This model is simple and easy to use. But there are many problems with this model. They can be found in [12]. The first problem is that under geometric Brownian motion, the spread can never be negative, yet negative spread do appear in most markets. Second, the lognormal also suggests that the volatility of the spread is positively related to its absolute value: the bigger spread is, the greater volatility gets. This is not supported by either market evidence or experience. Lastly, if one investor try to hedge a short position on a spread option, under this model, it only requires a single delta hedge strategy on the spread instead of two different strategies on each underlying asset. As pointed out in [12], this is clearly not sufficient.

Arithmetic Brownian Motion. In 1990, Wilcox [1] employed arithmetic Brownian motion to get a closed form formula for spread option. He assumes that the spread satisfies

$$dS_t = \alpha dt + \sigma dW_t \tag{3.6}$$

Then we have a closed form formula for the price of spread option:

$$V(t) = e^{-r(T-t)} \left[\left(S_t + \alpha(T-t) - K \right) N(w) + \sigma\sqrt{T-t}n(w) \right] \tag{3.7}$$

where

$$w = \frac{S_t + \alpha(T-t) - K}{\sigma\sqrt{T-t}} \tag{3.8}$$

and $N(w)$ and $n(w)$ represent the CDF and density function for normal distribution.

This model is simple but it has a fatal problem: it is not consistent with the no arbitrage condition. This point is mentioned in Poitras's paper in 1998 [2]. His argument is in the special case where both assets pay the same constant dividend yield δ, by treating the spread as one random variable, the price of the spread option $V(t, S_t)$ should satisfy

$$\frac{\partial V}{\partial t} = rV - (r-\delta)\frac{\partial V}{\partial x}x - \frac{1}{2}\frac{\partial^2 V}{\partial x^2}\sigma^2 \tag{3.9}$$

Now if we differentiate Wilcox's formula (3.7) with $\delta = 0$, we can see it does not satisfy the PDE because of the parameter α, which Poitras calls *arbitrary parameter*. To resolve this problem, in [2], Poitras imposed the condition $\alpha = (r-\delta)S_t$. Then the model of the spread is:

$$dS_t = (r-\delta)S_t dt + \sigma dW_t \tag{3.10}$$

this gives us a different formula:

$$V(t) = \left(S_t e^{-\delta(T-t)} - Ke^{-r(T-t)}\right) N(y) + Vn(y) \tag{3.11}$$

where

$$\begin{aligned} V &= \sigma\sqrt{\frac{e^{-2\delta(T-t)} - e^{-2r(T-t)}}{2(r-\delta)}} \\ y &= \frac{S_t e^{-\delta(T-t)} - Ke^{-r(T-t)}}{V} \end{aligned} \tag{3.12}$$

and $N(y)$ and $n(y)$ represent the CDF and density function for the normal distribution.

As mentioned in [12], all one-factor model suffers one problem is that it only provides single hedging strategy. The solution is that we need two-factor models.

3.1.2 Two-factor Models. In a two factor model, we consider the two underlying assets separately. The price process of each asset have its own SDE.

Arithmetic Brownian Motion. In [2], Poitras also proposed a two factor model. The two underlying assets follow

$$dS_{i,t} = (r - \delta_i)S_{i,t}dt + \sigma_i dW_{i,t} \tag{3.13}$$

for $i = 1, 2$. Where r is the interest rate, δ_i and σ_i are the dividend yield and volatility for each asset. Here $W_{i,t}$ are standard Brownian motion with correlation $dW_{1,t}dW_{2,t} = \rho dt$.

Under this model, the no-arbitrage pricing formula for spread call option is

$$V(t) = (S_{1,t}e^{-\delta_1}(T-t) - S_{2,t}e^{-\delta_2}(T-t) - Ke^{-r(T-t)})N(z) + \sigma n(z) \tag{3.14}$$

where

$$\begin{aligned}
\nu_{11} &= \sigma_1 \left(\frac{e^{-2\delta_1(T-t)} - e^{-2r(T-t)}}{2(r-\delta_1)} \right) \\
\nu_{22} &= \sigma_2 \left(\frac{e^{-2\delta_2(T-t)} - e^{-2r(T-t)}}{2(r-\delta_2)} \right) \\
\nu_{12} &= \rho\sigma_1\sigma_2 \left(\frac{e^{-(\delta_2+\delta_1)(T-t)} - e^{-2r(T-t)}}{2(r-\delta_1-\delta_2)} \right) \\
\sigma &= \sqrt{\sigma_1^2 + 2\rho\sigma_1\sigma_2 + \sigma_2^2} \\
z &= \frac{S_{1,t}e^{-\delta_1}(T-t) - S_{2,t}e^{-\delta_2}(T-t) - Ke^{-r(T-t)}}{\sigma}
\end{aligned} \tag{3.15}$$

Geometric Browinian Motion. A natural way for a two-factor geometric Browinian motion model would be assume the two assets satisfy

$$dS_{i,t} = \mu_i S_{i,t}dt + \sigma_i S_{i,t}dW_{i,t} \tag{3.16}$$

where $W_{i,t}$ are Brownian motion and $dW_{1,t}dW_{2,t} = \rho dt$. We know $S_{i,t}$ is lognormal and let $f(s_1|s_2)$ be the density function of $S_{2,t}$ given $S_{1,t}$, and $f(s_2)$ is the density function of $S_{2,t}$.

Then the price of a call option under risk-neutral measure is

$$\begin{aligned} V(t) &= \widetilde{E}[e^{-r(T-t)}(S_{1,T} - S_{2,T} - K)^+] \\ &= e^{-r(T-t)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (s_1 - s_2 - K)^+ f(s_1|s_2) f(s_2) ds_1 ds_2 \\ &= e^{-r(T-t)} \int_{-\infty}^{\infty} \int_{s_2+K}^{\infty} (s_1 - s_2 - K) f(s_1|s_2) ds_1 f(s_2) ds_2 \\ &= e^{-r(T-t)} \int_{-\infty}^{\infty} F(s_2) f(s_2) ds_2 \end{aligned} \tag{3.17}$$

where

$$F(s_2) = \int_{s_2+K}^{\infty} [s_1 - (s_2 + K)] f(s_1|s_2) ds_1 \tag{3.18}$$

We can consider $F(s_2)$ be the price of a call option with underlying asset $S_{1,t}$ with strike price $s_2 + K$, then we can use Black-Scholes-Merton formula. The solution for the geometric Browinian motion is given as:

$$S_{i,t} = S_{i,0} \exp\left\{(\mu_i - \frac{1}{2}\sigma_i)t + \sigma_i W_{i,t}\right\} \tag{3.19}$$

Now define two random variable $B_{1,t}$ and $B_{2,t}$ as

$$\begin{aligned} B_{1,t} &= W_{1,t} \\ B_{2,t} &= \frac{W_{2,t}}{\sqrt{1-\rho^2}} - \frac{\rho W_{1,t}}{\sqrt{1-\rho^2}} \end{aligned} \tag{3.20}$$

then we have following theorem

Theorem 3.1. *Let $W_{1,t}$ and $W_{2,t}$ be two Brownian motion with correlation $dW_{1,t}dW_{2,t} = \rho dt$. Let $B_{1,t}$ and $B_{2,t}$ be two random variable defined as in* (3.20)*, then they are independent Brownian motions.*

Proof. Since $W_{1,t}$ and $W_{2,t}$ are martingale, then both $B_{1,t}$ and $B_{2,t}$ are also martingale and have continuous path. We also have $B_{1,0} = B_{2,0} = 0$. It is easy to check that

$$
\begin{aligned}
dB_{1,t}dB_{1,t} &= dW_{1,t}dW_{1,t} = dt \\
dB_{2,t}dB_{2,t} &= \frac{dW_{2,t}dW_{2,t}}{1-\rho^2} - \frac{2\rho dW_{1,t}dW_{2,t}}{1-\rho^2} + \frac{\rho dW_{1,t}dW_{1,t}}{1-\rho^2} = \frac{1-2\rho^2+\rho}{1-\rho^2}dt = dt \\
dB_{2,t}dB_{2,t} &= \frac{\rho dt}{\sqrt{1-\rho^2}} - \frac{\rho dt}{\sqrt{1-\rho^2}} = 0
\end{aligned}
\tag{3.21}
$$

thus by the Levy's theorem, $B_{1,t}$, and $B_{2,t}$ are two independent Brownian motions. □

Now we can write (3.19) as

$$
\begin{aligned}
S_{1,t} &= S_{1,0} \exp\left\{(r - \frac{1}{2}\sigma_1^2)t + \sigma_1 B_{1,t}\right\} \\
S_{2,t} &= S_{2,0} \exp\left\{(r - \frac{1}{2}\sigma_2^2)t + \sigma_2(\rho B_{1,t} + \sqrt{1-\rho^2}B_{2,t})\right\}
\end{aligned}
\tag{3.22}
$$

Now let

$$
\begin{aligned}
\alpha_1 &= r - \frac{1}{2}\sigma_1^2 \\
\alpha_2 &= r - \frac{1}{2}\sigma_2^2 \\
\beta_1 &= \sigma_2\rho \\
\beta_1 &= \sigma_2\sqrt{1-\rho^2}
\end{aligned}
\tag{3.23}
$$

then the price of a European call options is

$$
\begin{aligned}
V(0) &= \widetilde{E}[e^{-rT}(S_{1,T} - S_{2,T} - K)^+] \\
&= \widetilde{E}[e^{-rT}(s_{1,0}\exp\{\alpha_1 T + \sigma_1 B_{1,T}\} - s_{2,0}\exp\{\alpha_2 T + \beta_1 B_{1,T} + \beta_2 B_{2,T}\} - K)^+] \\
&= e^{-rT}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(s_{1,0}\exp\{\alpha_1 T + \sigma_1 x_1\} - s_{2,0}\exp\{\alpha_2 T + \beta_1 x_1 + \beta_2 x_2\} - K)^+ f(x_1,x_2)dx_1dx_2
\end{aligned}
\tag{3.24}
$$

where x_1 and x_2 are two variables for the integration, and f is the joint density function of independent normally distributed random variables with mean zero and variance T. In this case we know the joint density function $f(x_1, x_2)$ is just the product of the density function for x_1 and x_2 since they are independent.

3.2 Existing Term-Structure Models

In this section, we will give a review for the existing term-structure models, specificly the Vasicek model. This model is one-factor, time homogeneous, meaning that there is only one random source for the interest rate and the rate only depends on constant coefficients. It is also called an *affine-yield* model, because the yield of a zero-coupon bond is assumed to be an affine function of the interest rate, i.e $P(t,T) = \exp\{-r(t)C(t,T) - A(t,T)\}$.

3.2.1 The Vasicek Model. Vasicek[13] assumed that the instantaneous spot rate follows the Ornstein-Uhlenbeck process:

$$dr(t) = \theta(\lambda - r(t))dt + \sigma dW_t \tag{3.25}$$

where $\theta,\ \lambda$ and σ are constants. The solution of this SDE is given by

$$r(t) = r(0)e^{-\theta t} + \lambda\left(1 - e^{-\theta t}\right) + \sigma \int_0^t e^{-\theta(t-u)} dW_u \tag{3.26}$$

so the rate $r(t)$ is normally distributed with mean and variance given by

$$\begin{aligned} E[r(t)] &= r(0)e^{-\theta t} + \lambda\left(1 - e^{-\theta t}\right) \\ Var[r(t)] &= \frac{\sigma^2}{2\lambda}\left(1 - e^{-2\theta t}\right) \end{aligned} \tag{3.27}$$

As you can see from the expectation of $r(t)$ as t goes to infinity, the expected rate tends to λ. So the parameter λ can be regarded as a long term average rate. So in this model $r(t)$ is *mean reverting*. The parameter θ is called the *mean revering speed*, it represents on average, how fast the interest rate changes with respect to the difference between the interest rate and the long term mean. Vasicek model is the first model to capture the mean reverting property of interest rate.

Since this is an affine-yield model, then the bond price is given as:

$$P(t,T) = \exp\{-r(t)C(t,T) - A(t,T)\} \tag{3.28}$$

where $A(t,T)$ and $C(t,T)$ are

$$C(t,T) = \frac{1}{\lambda}\left(1 - e^{-\lambda(T-t)}\right) \tag{3.29}$$

and

$$\begin{aligned} A(t,T) &= \int_t^T \lambda\theta C(s,T) - \frac{1}{2}\sigma^2 C^2(s,T)ds \\ &= \left(\theta - \frac{\sigma^2}{2\lambda^2}\right)(T - t - C(t,T)) + \frac{\sigma^2}{4\lambda}C^2 \end{aligned} \tag{3.30}$$

The proof and derivation of the bond price can be found in [6].

3.2.2 The Heath-Jarrow-morton Model.

The Heath-Jarrow-Morton (HJM) model models the yield curve in terms of forward rates. A forward rate $f(t,T)$ is a interest rate that you can lock at time t to borrow money at later time T. In [6], the author gives a really good example on forward rates, I will recite it here.

Let $\bar{T}$ be a fixed time horizon, all the bonds in this example will be mature at or before $\bar{T}$. For $0 \leq t \leq T \leq \bar{T}$, let $P(t,T)$ denote the price of a bond at time t that matures at time T with par value 1.

At time t, we can start so-called *forward investing* by setting up the following portfolio. Let $\delta > 0$ be small, the portfolio is as follows:

- Short one bond matures at time T.
- Long $\frac{P(t,T)}{P(t,T+\delta)}$ shares of bond matures at $T+\delta$.

you can see the cost of this portfolio is zero, since the short position generates income $P(t,T)$ and the long position costs $P(t,T)$. Now at time T, the investor will need to borrow one dollar to cover the bond he shorted, and at later time $T+\delta$, he receives $\frac{Pt,T}{P(t,T+\delta)}$ dollars from the long position. Now the value of the portfolios is

$$\frac{P(t,T)}{P(t,T+\delta)} - 1e^{g\delta} \tag{3.31}$$

where g is the interest rate to borrow at time T. Under the assumption that there is not arbitrage, we must have

$$\frac{P(t,T)}{P(t,T+\delta)} - 1e^{g\delta} = 0 \tag{3.32}$$

solve for g we get

$$g = -\frac{\ln P(t,T+\delta) - \ln P(t,T)}{\delta} \tag{3.33}$$

Note that g is known at time t,i.e. it is $\mathcal{F}_t$-measurable, thus an investor can "lock in" this rate at earlier time t. In fact this is the only rate an investor can lock in at time t to invest or borrow at time T without having arbitrage.

The *forward rate at time t for investing at time T* is defined as

$$f(t,T) = \lim_{\delta \to 0^+} -\frac{\ln P(t,T+\delta) - \ln P(t,T)}{\delta} = -\frac{\partial P(t,T)}{\partial T} \tag{3.34}$$

thus we have following relative between the bond price

$$P(t,T) = e^{-\int_t^T f(t,s)ds} \tag{3.35}$$

Note that the short interest rate at time t is $r(t) = f(t,t)$.

In the HJM model, assuming the initial forward curve $f(0,T)$ is known, the forward curve for a later time t satisfies the SDE

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW_t \tag{3.36}$$

or the integral form

$$f(t,T) = f(0,T) + \int_0^t \alpha(s,T)ds + \int_0^T \sigma(s,T)dW_s \tag{3.37}$$

where W_t is Brownian motion, and $\alpha(t,T)$ and $\sigma(t,T)$ is adapted to $\mathcal{F}_t$ and satisfies the conditions in order to have the solution exist. Thus the price of the bond $P(t,T)$ is given as

$$P(t,T) = \exp\left(\int_t^T f(0,s)ds + \int_t^T \int_0^t \alpha(u,s)duds + \int_t^T \int_0^t \sigma(u,s)dW_u ds\right) \tag{3.38}$$

by Fubini's theorem in real analysis and for Ito integral [14], we have

$$\begin{aligned} P(t,T) &= \exp\left(\int_t^T f(0,s)ds + \int_t^T \int_0^t \alpha(u,s)duds + \int_t^T \int_0^t \sigma(u,s)dW_u ds\right) \\ &= \exp\left(\int_t^T f(0,s)ds + \int_0^t \int_t^T \alpha(u,s)dsdu + \int_0^t \int_t^T \sigma(u,s)dsdW_u\right) \\ &= \exp\left(\int_t^T f(0,s)ds + \int_0^t \alpha^*(t,T)du + \int_0^t \Sigma(t,T)dW_u\right) \end{aligned} \tag{3.39}$$

where

$$\alpha^*(t,T) = \int_t^T \alpha(u,s)ds \quad \text{and} \quad \Sigma(t,T) = \int_t^T \sigma(u,s)ds \tag{3.40}$$

Recall that the short interest rate is

$$r(t) = f(t,t) = f(0,t) + \int_0^t \alpha(s,t)ds + \int_0^T \sigma(s,t)dW_s \tag{3.41}$$

then we can use this to find the dynamic of $f(t,T)$ under the risk-neutral measure. By Ito's formula and Gironove' theorem we have the SDE for the forward curve under risk-neutral measure as

$$df(t,T) = \sigma(t,T)\Sigma(t,T)dt + \sigma(t,T)d\widetilde{W}_t \tag{3.42}$$

and the bond price staisfies

$$dP(t,T) = r(t)P(t,T)dt - \Sigma(t,T)P(t,T)d\widetilde{W}_t \tag{3.43}$$

3.3 HJM and Short Rate Models

When pricing a derivative, the interest rate or short rate model is popular. Short rate models are HJM models, that's why we usually call the HJM model a framework. In this section, we will present the Vasicek model is a HJM model. To do so, we need to choose a volatility for the forward curve, so that the short rate model results from HJM model is exactly the same as the original model.

The Vasicek model can be written as

$$dr_t = \lambda(\mu - r(t))dt + \sigma r(t)d\widetilde{W}_t \tag{3.44}$$

Let $\alpha(r(t),t) = \lambda(\mu - r(t))$ and $\beta(r(t),t;\gamma) = \sigma$.

Then by the risk-neutral pricing formula, we know the price of a bond $P(t,T)$ is given as

$$P(t,T) = \widetilde{E}[\exp - \int_t^T -r(s)ds|\mathcal{F}_t] \tag{3.45}$$

then from (3.35) we have

$$\int_t^T f(t,s)ds = -\ln P(t,T) = -\ln \widetilde{E}\left[\exp - \int_t^T -r(s)ds|\mathcal{F}_t\right]. \tag{3.46}$$

By the Markov property of $r(t)$, we know there exists a function $g(x,t,T)$ such that

$$g(r(t),t,T) = -\ln \widetilde{E}\left[\exp - \int_t^T -r(s)ds|\mathcal{F}_t\right] \tag{3.47}$$

thus we have

$$f(t,T) = \frac{\partial g}{\partial T}. \tag{3.48}$$

Now we can apply Ito formula to $f(t,T)$ to get

$$df(t,T) = \left(\frac{\partial^2 g}{\partial T \partial t} + \alpha(r(t),t)\frac{\partial^2 g}{\partial T \partial x} + \frac{1}{2}\frac{\partial^3 g}{\partial T \partial x^2}\right)dt + \beta(r(t),t)\frac{\partial^2 g}{\partial T \partial x} d\widetilde{W}_t \tag{3.49}$$

thus we should choose the forward volatility to be

$$\sigma(t,T) = \beta(r(t),t)\frac{\partial^2 g}{\partial T \partial x}. \tag{3.50}$$

From here we know the volatility for the bond price process is given

$$\Sigma(t,T) = \int_t^T \sigma(u,s)ds = \beta(r(t),t)\int_t^T \frac{\partial^2 g}{\partial T \partial x}(r(t),t,s)ds = \beta(r(t),t)\frac{\partial g}{\partial x} \tag{3.51}$$

In the Vasicek model, since the bond price $P(t,T)$ is given in (3.28), then we know

$$\frac{\partial^2 g}{\partial T \partial x}(x,t,T) = \frac{\partial C}{\partial T} \tag{3.52}$$

and

$$\frac{\partial g}{\partial x} = C(t,T) \tag{3.53}$$

From (3.29) and (3.30), we can get the forward volatility for Vasicek Model is

$$\sigma^F(t,T) = \sigma \exp\{-\lambda(T-t)\} \tag{3.54}$$

and volatility for the bond price is

$$\Sigma(t,T) = \int_t^T \sigma \exp\{-\lambda(T-s)\}ds = \frac{\sigma}{\lambda}\left(e^{-\lambda(T-t)} - 1\right) \tag{3.55}$$

Chapter 4. Mathematical Analysis of the Model

In this chapter, we will introduce the pricing model with stochastic interest rate. Following the work of Black, Scholes, Merton and Shreve, we will use the Forward measure to derive a partial different equation for the price of a call spread option. The PDE has a free boundary. In the first section, we will review numeraire and change of measure. In section 2, we will see an example of using EMM to derive a partial differential equation under a constant interest rate. Following the example, we will derive the partial differential equation for the stochastic interest rate model.

4.1 Numeraire

A numeraire is an asset of a strictly positive price process. We can see it as a reference of the value of other assets. Money or dollars is one of the most common used numeraire. It always has strictly positive value, and we use it to reference how much an asset is worth. In real market, most assets and securities are quoted in dollars. But the currencies have time value which implies that the prices in terms of dollar or Yuan may not be consistent in time. In [15], Vecer pointed out that since a dollar today is worth more than a dollar tomorrow, the price of an asset in terms of dollar will have an upward drift component that corresponds to the loss of value of the dollar. Therefore the price process in dollars is not always martingale. Thus we need to find a new reference or numeraire in option pricing in order to get an arbitrage-free price.

The money-market account has long served as a numeraire in lots of option pricing literature, but by no means the only choice. In later chapter, we shall use a zero coupon bond as a numeraire to simplify derivative valuation. For each numeraire, we associate an equivalent martingale measure (EMM). This measure is equivalent to the measure we used to describe the price process before change of numeraire, and under EMM the price process "discounted" by the numeraire is a martingale.

Consider a portfolio X that contains assets $S_0, \cdots S_d$, with S_0 being the money-market. Let $\Delta = (\Delta_0, \ldots, \Delta_d)$ be the positions in each asset, then the value of the portfolio is given by

$$X(\Delta, t) = \sum_{i=1}^{d} \Delta_i S_i(t) \tag{4.1}$$

In order to have the ability to continuously adjust position in each asset without withdrawals and injections of funds, we require this portfolio to be *self-financing*, which is expressed as

$$dX(\Delta, t) = \sum_{i=1}^{d} \Delta_i dS_i(t) \tag{4.2}$$

Let N_t be a numeraire, and let S_i^N denote the new price based on N_t, $S_i^N = S_i/N$, then as proved in [16] we have the following Numeraire Invariance Theorem

Theorem 4.1. *Self-financing portfolios reman self-financing after a numeraire change.*

Let $\mathbb{P}^N$ denote the EMM associated with N_t, then since $S_i^N(t)$ is martingale for $i = 0, \cdots, d$, then the price of portfolio based on N_t denoted by X^N is also an martingale, i.e. for $t < T$, we have

$$X^N(\Delta, t) = \widetilde{E}^N[X^N(\Delta, T)|\mathcal{F}_t] \tag{4.3}$$

thus we can have the formula

$$\begin{aligned} X(\Delta, t) &= N_t \cdot X^N(\Delta, t) \\ &= N_t \cdot \widetilde{E}^N[X^N(\Delta, T)|\mathcal{F}_t] \\ &= N_t \cdot \widetilde{E}^N\left[\frac{X(\Delta, T)}{N_T}|\mathcal{F}_t\right] \\ &= \widetilde{E}^N\left[\frac{N_t}{N_T} \cdot X(\Delta, T)|\mathcal{F}_t\right] \end{aligned} \tag{4.4}$$

When we price a derivative,denote its value $V(t)$, if we can create a portfolio $X(\Delta, t)$ that replicates the value of the derivative at its expiration, then the value of the derivative at any time $t < T$ has to equal the value of the portfolio $X(\Delta, t)$, otherwise there will be an arbitrage. If $V(t) > X(\Delta, t)$ for some $t < T$, we short the derivative and use the fund to create the portfolio, and put the extra fund in the bank. At the expiration, the portfolio will offset the derivative, and our risk free profit is the extra fund plus interest. If the derivative is worth less than the portfolio, we do the opposite. Thus $X(\Delta, t)$ is the arbitrage free price and $V(t) = X(\Delta, t)$, then (4.4) can be reduced to the following pricing formula

$$V(t) = \widetilde{E}^N \left[\frac{N_t}{N_T} \cdot V(T) | \mathcal{F}_t \right] \tag{4.5}$$

Now we can use (4.5) to price its value at any time $t < T$. In a market such that every derivative can be hedged, we call the market *complete*. The following theorem in [17] and [18] answered the question when the market is complete under a EMM,

Theorem 4.2. ***Fundamental Theorem of Asset Pricing*** *Given a Numraire N, the following statements hold*

(i) A market is complete if and only if every derivative can be hedged

(ii) If there exists an EMM associated with N, then there is no arbitrage opportunity

(iii) The EMM is unique if and only if the market is complete.

We can price a derivative with two different numeraire N_t and M_t, the price $V(t)$ should be the same under these two numeraire, otherwise there would be arbitrage. Thus we have the following change of numeraire formula,

$$\widetilde{E}^N \left[\frac{N_t}{N_T} \cdot V(T) | \mathcal{F}_t \right] = \widetilde{E}^M \left[\frac{M_t}{M_T} \cdot V(T) | \mathcal{F}_t \right] \tag{4.6}$$

4.2 Martingale Approaches

In this section, we will use the constant interest rate model as an example, to show how to use an EMM to derive a partial differential equation for the price of a spread option. In this example the numeraire is the money market account. Let's consider the case for a European style spread call, the two assets with strike price K and maturity T, and price process of the two underlying assets satisfying the following SDEs

$$
\begin{aligned}
dS_{1,t} &= \alpha_1(t)S_{1,t}dt + \sigma_t S_{1,t}dW_{1,t} \\
dS_{2,t} &= \alpha_1(t)S_{2,t}dt + \sigma_t S_{2,t}dW_{2,t}
\end{aligned}
\tag{4.7}
$$

on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and where $\alpha_i(t)$ and $\sigma_i(t)$ are adapted processes. Let $W(t) = (W_1(t), W_1(t))$ be a two dimensional Brownian motion and r be the risk-free rate and here it is assumed to be a constant. Under the measure $\mathbb{P}$, the price process is not martingale, thus we need to find a new measure such that the discounted price process is martingale. This can be done with Girsanov theorem.

Following the literature in option pricing, let's use the money-market account as a numeraire. If we invest one dollar in the money-market at time 0, then the value of that dollar at time t is e^{rt}, where r is the risk-free rate and here is assumed to be a constant. Then the value of the two assets' price, discounted by the money-market account, is $e^{-rt}S_{i,t}$ for $i = 1, 2$. We want to find a new measure, such that this price process is martingale.

Define $\theta(t) = (\theta_1(t), \theta_2(t))$ as

$$
\theta_i(t) = \frac{\alpha_i(t) - r}{\sigma_i(t)} \tag{4.8}
$$

for $i = 1, 2$, and define two processes as:

$$
\begin{aligned}
Z(t) &= \exp\left\{ -\int_0^t \theta(u) \cdot dW(u) - \frac{1}{2}\int_0^t \|\theta(u)\|^2 du \right\}, \\
\widetilde{W}(t) &= W(t) + \int_0^t \theta(u) \cdot du
\end{aligned}
\tag{4.9}
$$

and a measure $\widetilde{\mathbb{P}}$ as

$$
\widetilde{\mathbb{P}}(A) = \int_A Z(T)dP \quad \text{for all } A \in \mathcal{F} \tag{4.10}
$$

then by the Girsanov theorem [6], the processes $\widetilde{W}(t)$ is a two dimensional Brownian motion.

Under the measure $\widetilde{\mathbb{P}}$, the discounted pricing process $e^{-rt}S_{i,t}$ is a martingale, for $i = 1, 2$.

Moreover, we can use Ito formula to get the new SDE for assets' pricing process

$$\begin{aligned} dS_{1,t} &= rS_{1,t}dt + \sigma_t S_{1,t} d\widetilde{W}_{1,t} \\ dS_{2,t} &= rS_{2,t}dt + \sigma_t S_{2,t} d\widetilde{W}_{2,t} \end{aligned} \tag{4.11}$$

As you can see under $\widetilde{\mathbb{P}}$, the expected return for the asset is the risk-free rate, this means that all the investors do not need to extract premium to invest in risky asset, thus they are risk-neutral. We call the measure $\widetilde{\mathbb{P}}$ the risk-neutral measure.

Let $V(t)$ be the price of the call at time t, then $V(T)$ would be the payoff of this option. Under the risk-neutral measure, we have following pricing formula:

$$V(t) = \tilde{E}[e^{-r(T-t)}V(T)|\mathcal{F}_t] \tag{4.12}$$

then by the iterating condition of the conditional expectation [19], we have for $s \leq t \leq T$,

$$\begin{aligned} \widetilde{E}[e^{-r(t)}V(t)|\mathcal{F}_s] &= \widetilde{E}[\widetilde{E}[e^{-r(T)}V(T)|\mathcal{F}_t]|\mathcal{F}_s] \\ &= \widetilde{E}[e^{-r(T)}V(T)|\mathcal{F}_s] \\ &= e^{-rs}V(s) \end{aligned} \tag{4.13}$$

thus the process $e^{-r(t)}V(t)$ is a martingale. By the Feynman-Kac theorem and the Markove property of solutions of stochastic equations, there exists a function $c(t, x, y)$ such that

$$c(t, S_{1,t}, S_{2,t}) = V(t) = \tilde{E}[e^{-r(T-t)}V(T)|\mathcal{F}_t] \tag{4.14}$$

for the details of Feynman-Kac theorem and the Markove property, please read [6] and [5].

Combine (4.13) and (4.14), we know that $e^{-r(t)}c(t, S_{1,t}, S_{2,t})$ is a martingale. Then by

Ito's formula we have

$$
\begin{aligned}
d(e^{-r(t)}c(t,S_{1,t},S_{2,t})) = e^{-r(t)}[&-rc(t,S_{1,t},S_{2,t})dt + c_t(t,S_{1,t},S_{2,t})dt + \\
&c_{S_{1,t}}(t,S_{1,t},S_{2,t})dS_{1,t} + c_{S_{2,t}}(t,S_{1,t},S_{2,t})dS_{2,t} + \\
&c_{S_{1,t}S_{1,t}}(t,S_{1,t},S_{2,t})dS_{1,t}dS_{2,t} + \\
&\frac{1}{2}c_{S_{2,t}S_{2,t}}(t,S_{1,t},S_{2,t})(dS_{2,t})^2 + \\
&\frac{1}{2}c_{S_{1,t}S_{1,t}}(t,S_{1,t},S_{2,t})(dS_{1,t})^2]
\end{aligned}
\tag{4.15}
$$

now plug (4.11) into (4.15), the drift term of $d(e^{-r(t)}c(t,S_{1,t},S_{2,t}))$ is

$$
e^{-rt}\left(-rc + c_t + rS_{1,t}c_{S_{1,t}} + rS_{2,t}c_{S_{2,t}} + \sigma_1\sigma_2 S_{1,t}S_{2,t}c_{S_{1,t}S_{2,t}} + \frac{1}{2}\sigma_1^2 S_{1,t}^2 c_{S_{1,t}S_{1,t}} + \frac{1}{2}\sigma_2^2 S_{2,t}^2 c_{S_{2,t}S_{2,t}}\right) \tag{4.16}
$$

since $e^{-r(t)}c(t,S_{1,t},S_{2,t})$ is martingale, then its drift term is zero, now replace $S_{1,t}$ and $S_{2,t}$ by the variables, x, y we have following PDE for the price of a spread option

$$
-rc + c_t + rxc_x + ryc_y + \sigma_1\sigma_2 xyc_{xy} + \frac{1}{2}\sigma_1^2 x^2 c_{xx} + \frac{1}{2}\sigma_2^2 y^2 c_{yy} = 0 \tag{4.17}
$$

with the terminal condition

$$
c(T,x,y) = (x-y-K)^+ \tag{4.18}
$$

and boundary conditions

$$
\begin{aligned}
&c(t,0,y) = 0 \\
&c(t,x,0) = c_{BS}(t,x) \\
&\lim_{x\to\infty} c(t,x,y) = x \\
&\lim_{y\to\infty} c(t,x,y) = 0
\end{aligned}
\tag{4.19}
$$

where $c_{BS}(t,x)$ is the price of a call option with underlying asset $S_{1,t}$ and same strike price.

4.3 Martingale Approach when Interest Rate is Stochastic

Assume that under the risk-natural measure, the price process for the two underlying assets satisfy the following stochastic differential equation:

$$\begin{aligned}\frac{dS_{1,t}}{S_{1,t}} &= r(t)dt + \sigma_1 d\widetilde{W}_{1,t} \\ \frac{dS_{2,t}}{S_{2,t}} &= r(t)dt + \sigma_2 d\widetilde{W}_{2,t}\end{aligned} \tag{4.20}$$

where σ_1 and σ_2 are volatility for the two underlying assets, which are assumed to be constant here, and $r(t)$ is the interest rate, which is assumed to satisfy

$$dr_t = \lambda(\mu - r_t)dt + \sigma_3 r_t^{\gamma} d\widetilde{W}_{3,t} \tag{4.21}$$

where λ is the speed of mean reversion and μ is the long term mean, σ_3 is the volatility for the interest rate. Here all three parameters are assumed to be constant. In the next chapter, we will review different methods to estimate these parameters from market data.

The Brownian motion $\widetilde{W}_{1,t}$, $\widetilde{W}_{2,t}$, and $\widetilde{W}_{3,t}$ are assumed to have following correlation:

$$\begin{aligned}d\widetilde{W}_{1,t}d\widetilde{W}_{2,t} &= \rho_1 dt \\ d\widetilde{W}_{1,t}d\widetilde{W}_{3,t} &= \rho_2 dt \\ d\widetilde{W}_{2,t}d\widetilde{W}_{3,t} &= \rho_3 dt\end{aligned} \tag{4.22}$$

Let $V(t)$ denote the value of a spread call option for the two underlying asset, with strike price K and expiration date T at time $t \leq T$, then

$$V(T) = \max\{S_{1,T} - S_{2,T} - K, 0\} \tag{4.23}$$

and with the risk-neutral formula, we have

$$V(t) = \frac{\widetilde{E}[D(T)V(T)|\mathcal{F}_t]}{D(t)} \tag{4.24}$$

where $D(t)$ is the discount factor:

$$D(t) = \exp\left\{-\int_0^t r(u)du\right\} \tag{4.25}$$

We will be following the martingale approach in the previous section to get a partial differential equation for the price of a call spread option under stochastic interest rate.

4.3.1 Forward Measure. The risk-neutral measure is one of the most used equivalent martingale measures in mathematical finance, the numeraire is the money market. Forward measure is another equivalent martingale measure, and it is used when the interest rate is stochastic. In Shreve's book [6], he used forward measure to derive a formula for the price of an option with stochastic interest rate. In this section, I will follow his work, and derive a partial differential equation for the price of a spread option with stochastic interest rate. The numeraire for forward measure is the bond with par value 1 with expiration time T, namely $P(t,T)$.

From the change of numeraire formula (4.6), we have:

$$\widetilde{E}^F[X|\mathcal{F}(s)] = \frac{\widetilde{E}[XD(t)P(t,T)|\mathcal{F}_s]}{D(s)P(s,T)} \tag{4.26}$$

Let $\widetilde{E}^F[\cdot]$ denote the expectation of a random value under the forward measure. Then by the change of measure formula (4.26) we have

$$\begin{aligned}V(t) &= \frac{\widetilde{E}[D(T)V(T)|\mathcal{F}_t]}{D(t)} \\ &= P(t,T)\frac{\widetilde{E}[D(T)V(T)P(T,T)|\mathcal{F}_t]}{D(t)P(t,T)} \\ &= P(t,T)\widetilde{E}^F[V(T)|\mathcal{F}_t] \\ &= P(t,T)\widetilde{E}^F[\max\{S_{1,T}-S_{2,T}-K,0\}|\mathcal{F}_t]\end{aligned} \tag{4.27}$$

since $(S_{1,t},S_{2,t},r(t))$ is Markov process, then there exists a Borel measurable function $f(t,x,y,z)$ such that

$$f(t,S_{1,t},S_{2,t},r(t)) = \widetilde{E}^F[\max\{S_{1,T}-S_{2,T}-K,0\}|\mathcal{F}_t] \tag{4.28}$$

then we have

$$V(t) = P(t,T)f(S_{1,t},S_{2,t},t) \tag{4.29}$$

and

$$\frac{V(t)}{P(t,T)} = f(S_{1,t},S_{2,t},t) \tag{4.30}$$

so if we know the function f, then multiply it by the value of a bond with par value 1 and maturity T, we should get the price of this option. In order to get a PDE, we still need to know what process $S_{1,t}, S_{2,t}$ and $P(t,T)$ satisfy under forward measure.

Let $\widetilde{B}_{1,t},\ \widetilde{B}_{2,t},\ \widetilde{B}_{3,t}$ be three random process such that

$$\begin{aligned}d\widetilde{W}_{1,t} &= d\widetilde{B}_{1,t} \\ d\widetilde{W}_{2,t} &= \rho_1 d\widetilde{B}_{1,t} + \sqrt{1-\rho_1^2}\,d\widetilde{B}_{2,t} \\ d\widetilde{W}_{3,t} &= \rho_2 d\widetilde{B}_{1,t} + \frac{\rho_3-\rho_1\rho_2}{\sqrt{1-\rho_1^2}}d\widetilde{B}_{2,t} + \sqrt{1-\rho_2^2-\frac{(\rho_3-\rho_1\rho_2)^2}{1-\rho_1^2}}\,d\widetilde{B}_{3,t}\end{aligned} \tag{4.31}$$

then by Levy's theorem, they are independent Brownian motion under the risk-neutral measure.

For simplicity, let

$$A_1 = \frac{\rho_3 - \rho_1\rho_2}{\sqrt{1-\rho_1^2}} \quad \text{and} \quad A_2 = \sqrt{1-\rho_2^2 - \frac{(\rho_3-\rho_1\rho_2)^2}{1-\rho_1^2}} \tag{4.32}$$

Since bond is a tradeable asset, then under the risk-neutral measure, the price process $P(t,T)$ must have the following SDE:

$$\frac{dP(t,T)}{P(t,T)} = r(t)dt + \sigma_P d\widetilde{W}_{3,t} \tag{4.33}$$

where σ_P is the volatility of $P(t,T)$, it is given in (3.55) for the Vasicek model.

The SDE for $S_{1,t},\ S_{2,t},\ r(t)$ and $P(t,T)$ are

$$\begin{aligned}
\frac{dS_{1,t}}{S_{1,t}} &= r(t)dt + \sigma_1 d\widetilde{B}_{1,t} \\
\frac{dS_{2,t}}{S_{2,t}} &= r(t)dt + \sigma_2\rho_1 d\widetilde{B}_{1,t} + \sigma_2\sqrt{1-\rho_1^2}\,d\widetilde{B}_{2,t} \\
dr(t) &= \lambda(\mu - r(t))dt + \sigma_3\rho_2 d\widetilde{B}_{1,t} + \sigma_3 A_1 d\widetilde{B}_{2,t} + \sigma_3 A_2 d\widetilde{B}_{3,t} \\
\frac{dP_{t,T}}{P_{t,T}} &= r(t)dt + \sigma_P\rho_2 d\widetilde{B}_{1,t} + \sigma_P A_1 d\widetilde{B}_{2,t} + \sigma_P A_2 d\widetilde{B}_{3,t}
\end{aligned} \tag{4.34}$$

Now let's change the measure by letting

$$\theta_t = (-\sigma_P\rho_2, -\sigma_P A_1, -\sigma_P A_2) \tag{4.35}$$

we will get new independent Brownian motions as

$$\begin{aligned}
d\widetilde{B}^F_{1,t} &= -\sigma_P\rho_2 dt + d\widetilde{B}_{1,t} \\
d\widetilde{B}^F_{2,t} &= -\sigma_P A_1 dt + d\widetilde{B}_{2,t} \\
d\widetilde{B}^F_{3,t} &= -\sigma_P A_2 dt + d\widetilde{B}_{3,t}
\end{aligned} \tag{4.36}$$

and new SDEs for $S_{1,t},\ S_{2,t}$ and $r(t)$

$$\begin{aligned}
\frac{dS_{1,t}}{S_{1,t}} &= (r(t) + \sigma_1\sigma_P\rho_2)dt + \sigma_1 d\widetilde{B}^F_{1,t} \\
\frac{dS_{2,t}}{S_{2,t}} &= \left(r(t) + \sigma_2\sigma_P\rho_1\rho_2 + \sigma_2\sigma_P A_1\sqrt{1-\rho_1^2}\right)dt + \sigma_2\rho_1 d\widetilde{B}^F_{1,t} + \sigma_2\sqrt{1-\rho_1^2}d\widetilde{B}^F_{2,t} \\
dr(t) &= \left[\lambda(\mu - r(t)) + \sigma_3\sigma_B\rho_2^2 + \sigma_3\sigma_B A_1^2 + \sigma_3\sigma_B A_2^2\right]dt + \sigma_3\rho_2 d\widetilde{B}^F_{1,t} + \sigma_3 A_1 d\widetilde{B}^F_{2,t} + \sigma_3 A_2 d\widetilde{B}^F_{3,t}
\end{aligned} \tag{4.37}$$

Recall the pricing formula under forward measure

$$V(t) = P(t,T)\widetilde{E}^F[\max\{S_{1,T} - S_{2,T} - K, 0\}|\mathcal{F}_t] \tag{4.38}$$

and the Borel measurable function $f(t,x,y,z)$, such that

$$f(t, S_{1,t}, S_{2,t}, r(t)) = \widetilde{E}^F[\max\{S_{1,T} - S_{2,T} - K, 0\}|\mathcal{F}_t] = \frac{V(t)}{P(t,T)} \tag{4.39}$$

Applying Ito formula we have

$$\begin{aligned}
d\left(\frac{V(t)}{P(t,T)}\right) = d(t, f(S_{1,t}, S_{2,t}, r(t))) =& f_t dt + f_{S_{1,t}} dS_{1,t} + f_{S_{2,t}} dS_{2,t} + f_{r(t)} dr(t) + \\
&+ \frac{1}{2} f_{S_{1,t}S_{1,t}}(dS_{1,t})^2 + \frac{1}{2} f_{S_{2,t}S_{2,t}}(dS_{2,t})^2 + \frac{1}{2} f_{r(t)r(t)}(dr(t))^2 \\
&+ f_{S_{1,t}S_{2,t}} dS_{1,t}dS_{2,t} + f_{S_{1,t}r(t)} dS_{1,t}dr(t) + f_{S_{2,t}r(t)} dS_{2,t}dr(t)
\end{aligned} \tag{4.40}$$

where

$$
\begin{aligned}
&f_{S_{1,t}}dS_{1,t} = f_{S_{1,t}}S_{1,t}(r(t)+\sigma_1\sigma_P)dt + M_1\\
&f_{S_{2,t}}dS_{2,t} = f_{S_{2,t}}S_{2,t}(r(t)+\sigma_2\sigma_P\rho_1\rho_2+\sigma_2\sigma_P A_1\sqrt{1-\rho_1^2})dt + M_2\\
&f_{r(t)}dr(t) = f_{r(t)}\left(\lambda(\mu-r(t))+\sigma_3\sigma_B\rho_2^2+\sigma_3A_1^2\sigma_B-\sigma_3\sigma_BA_2^2\right)dt + M_3\\
&f_{S_{1,t}S_{1,t}}(dS_{1,t})^2 = f_{S_{1,t},S_{1,t}}\sigma_1^2S_{1,t}^2dt\\
&f_{S_{2,t}S_{2,t}}(dS_{2,t})^2 = f_{S_{2,t},S_{2,t}}\sigma_1^2S_{2,t}^2dt\\
&f_{r(t)r(t)}(dr(t))^2 = f_{r(t)r(t)}\left(\sigma_3^2\rho_2^2+\sigma_3^2A_1^2+\sigma_3^3A_2^2\right)dt\\
&f_{S_{1,t}S_{2,t}}dS_{1,t}dS_{2,t} = f_{S_{1,t},S_{2,t}}\sigma_1\sigma_2\rho_1dt\\
&f_{S_{1,t}r(t)}dS_{1,t}r(t) = f_{S_{1,t}r(t)}\sigma_1\sigma_3\rho_2dt\\
&f_{S_{2,t}r(t)}dS_{2,t}dr(t) = f_{S_{2,t}r(t)}\left(\sigma_2\sigma_3\rho_1\rho_2+\sigma_2\sigma_3\sqrt{1-\rho_1^2}A_1\right)dt
\end{aligned}
\tag{4.41}
$$

where $M_1,\ M_2$ and M_3 are the diffusion part. Let

$$
\begin{aligned}
D_1 &= \sigma_3\sigma_B\rho_2^2+\sigma_3A_1^2\sigma_B-\sigma_3\sigma_BA_2^2\\
D_2 &= \sigma_2\sigma_P\rho_1\rho_2+\sigma_2\sigma_PA_1\sqrt{1-\rho_1^2}\\
D_3 &= \sigma_3^2\rho_2^2+\sigma_3^2A_1^2+\sigma_3^3A_2^2\\
D_4 &= (\sigma_2\sigma_3\rho_1\rho_2+\sigma_2\sigma_3\sqrt{1-\rho_1^2}A_1
\end{aligned}
\tag{4.42}
$$

then the drift term of $d(\frac{V(t)}{P(t,T)})$ is

$$
\begin{aligned}
&f_t + f_{r(t)}(\lambda(\mu-r(t))+D_1) + f_{S_{1,t}}S_{1,t}(r(t)\sigma_1\sigma_B\rho_2) + f_{S_{2,t}}S_{2,t}(r(t)+D_2)\\
&+\frac{1}{2}f_{S_{1,t},S_{1,t}}\sigma_1^2S_{1,t}^2+\frac{1}{2}f_{S_{2,t},S_{2,t}}\sigma_1^2S_{2,t}^2+\frac{1}{2}f_{r(t)r(t)}D_3\\
&+f_{S_{1,t},S_{2,t}}S_{1,t}S_{2,t}\sigma_1\sigma_2\rho_1+f_{S_{1,t}r(t)}S_{1,t}\sigma_1\sigma_3\rho_2+f_{S_{2,t}r(t)}D_4
\end{aligned}
\tag{4.43}
$$

Note that under forward measure, the "discounted" price is a martingale, then the drift term of $d(\frac{V(t)}{P(t,T)})$ should be zero. Replace $S_{1,t}, S_{2,t}$ and $r(t)$ by x, y, z in (4.43) and set it equal to zero we have the following PDE

$$
\begin{aligned}
0 =& f_t + f_z(\lambda(\mu - z) + D_1) + f_x x z \sigma_1 \sigma_B \rho_2 + f_y y(z + D_2) + \frac{1}{2} f_{xx} \sigma_1^2 x^2 + \frac{1}{2} f_{yy} \sigma_1^2 y^2 + \\
& \frac{1}{2} f_{zz} D_3 + f_{xy} x y \sigma_1 \sigma_2 \rho_1 + f_{xz} \sigma_1 \sigma_3 \rho_2 + f_{yz} D_4
\end{aligned}
\tag{4.44}
$$

with the terminal condition

$$
f(T, x, y, z) = \max(x - y - K, 0); \tag{4.45}
$$

the boundary conditions are

$$
\begin{aligned}
& f(t, 0, y, z) = 0 \\
& f(t, x, 0, z) = BS(t) \\
& f(t, x, y, 0) = g(x, y) \\
& \lim_{x \to \infty} f(t, x, y, z) = x \\
& \lim_{y \to \infty} f(t, x, y, z) = 0 \\
& \lim_{z \to \infty} f(t, x, y, z) = 0
\end{aligned}
\tag{4.46}
$$

$BS(t)$ is the price of an option with stochastic interest rate and the underlying asset is S_1 and the same strike price. There is a closed formula provided in [6], it is given by

$$
BS(t) = S(t)N(d_+(t)) - KP(t, T)N(d_-(t)) \tag{4.47}
$$

where $P(t, T)$ is the price of a zero-coupon bond with face value one, expires at time T and the processes $d_\pm(t)$ are given

$$
d_\pm(t) = \frac{1}{\sigma\sqrt{T - t}} \left[\log \frac{For_S(t, T)}{K} \pm \frac{1}{2} \sigma^2 (T - t) \right] \tag{4.48}
$$

and $For_S(t, T)$ is the forward price on the asset S.

Note this PDE has a free boundary, when the interest rate is zero, the price of the option today should equal the payoff at expiration T, which is unknown at time t. So in order to calculate the price, we can use the Monte Carlo methods.

Chapter 5. Numerical Computations

The partial differential equation (4.44) in the previous section 4.3 is hard to solve analytically and even numerically because of the free boundary condition. In this Chapter, we will study the price of an spread option numerically using Monte Carlo simulation. In section 1, we introduce different methods to estimate the parameters in the Vasicek model. In section 2, we will show the numerical result on how different values of the parameters in the Vasicek model will change the option price.

5.1 Parameters Estimate

In this section, I will present the Generalized Methods of Moments (GMM) estimation procedure due to Hansen [20]. The key advantage of GMM is that it requires specification only of certain moments not the density. We will also present that the two commonly used estimation methods Ordinary Least Squares (OLS) and Maximum Likelihood Estimation (MLE) can be viewed as a special case of GMM.

5.1.1 Generalized Methods of Moments. Hansen's formulation of the estimation problem is as follows. Let x_t be a $k \times 1$ vector of observations at time t, and x_t is assumed to be stationary. Let θ denote the $a \times 1$ vector of unknown parameters are being estimated. Let $h(\cdot, \cdot)$ be a $r \times 1$ vector-valued function that maps from $\mathbb{R}^a \times \mathbb{R}^k$ to $\mathbb{R}^r$. Let θ_0 denote the true value of θ and suppose that the function h satisfies the *orthogonality conditions*[21]

$$E[h(\theta_0, w_t)] = 0 \tag{5.1}$$

Let $\{x_t \in \mathbb{R}^k, t = 1, 2, \ldots, T\}$ be the collection of observations of size T, and Y_T be a $Tk \times 1$ vector with all observations $Y_T := (x_1, x_2, \cdots, x_T)$. Define a $r \times 1$ vector-valued function $g(\cdot) : \mathbb{R}^a \to \mathbb{R}^r$ as

$$g(\theta; Y_T) = \frac{1}{T} \sum_{t=1}^{T} h(\theta, x_t) \tag{5.2}$$

Note that the vector Y_T is a parameter for g, not a variable. The idea behind the GMM is that to find the value of θ that makes the sample moment $g(\theta, Y_T)$ to be as close as possible to the true moment zero, let $\hat{\theta}$ denote this estimator. This is the same to find the vector θ that minimized the quadratic

$$Q(\theta) := g(\theta; Y_T)' W_T g(\theta, Y_T) \tag{5.3}$$

Here the prime denotes the transpose of the vector. The $r \times r$ matrix W_T is a positive semi-definite matrix called the weighting matrix.

If the number of the unknown parameters a is equal to the number of orthogonality conditions r, then we call this system exactly identified. Then the sample moment function can be minimized by solving the equation

$$g(\theta, Y_t) = 0 \tag{5.4}$$

If $r > a$, then we call this system over-identified, and (5.4) cannot hold exactly. That is the reason we are using the weighting matrix W_T. How close is the *ith* condition to zero totally depends on how much of the weight is given to this condition.

The estimator $\widehat{\theta}$ depends on the weighting matrix W_T.The optimal weighting matrix is given by S^{-1}, and S is the asymptotic variance of the sample mean of $h(\theta_0, w_t)$, it is given in[20] and [21]:

$$S = \lim_{T \to \infty} T \cdot E[g(\theta_0, Y_T)' g(\theta_0, Y_T)] \tag{5.5}$$

then the GMM estimator $\hat{\theta}$ is obtained by minimizing the quadratic

$$Q(\theta) := g(\theta; Y_T)' S^{-1} g(\theta, Y_T) \tag{5.6}$$

Note that the description of the optimal weighting matrix is circular, we need to know S to estimate θ, but we need to know the true value θ_0 to get S, BUT if we know θ_0 at the beginning, there is no point to do all this estimation. To solve this issue, the strategy is as follows: we get an initial estimation $\hat{\theta}^0$ by using the identity matrix as the weighting matrix, then we use $\hat{\theta}^0$ to estimate S by using S_T

$$S_T = \lim_{T\to\infty} T \cdot E[g(\hat{\theta}, Y_T)' g(\hat{\theta}, Y_T)] \tag{5.7}$$

then (5.6) is minimized again with $S = S_T$ to get a new estimator $\hat{\theta}^1$. We keep doing this iteration until $\|\hat{\theta}^i - \hat{\theta}^{i+1}\|$ is less than some tolerance.

We have to make certain assumptions on orthogonality conditions h, the weighting matrix W_T and the parameter space Θ to obtain the asymptotic distribution of the GMM estimator $\hat{\theta}$, we will not discuss these assumptions here and interested readers are refered to [20] and [22]. For any weighting matrix, the estimator is consistent, and further assumption such as

$$\frac{1}{\sqrt{T}} \sum_{t=0}^{T} g(\theta, x_t) \longrightarrow \mathcal{N}(0, S) \tag{5.8}$$

the convergence is in distribution. The the GMM estimator has the following standard error:

$$\sqrt{T}(\hat{\theta} - \theta_0) \sim \mathcal{N}(0, V_T) \tag{5.9}$$

where

$$V_T = (D_T S_T^{-1} D_T')^{-1} \tag{5.10}$$

and D_T' is the Jacobian matrix of $g(\theta; Y_T)$ with respect to θ. We will show below that the OLS and MLE are special cases of GMM.

5.1.2 Ordinary Least Squares. OLS is a method for estimating unknown parameters by linear regression. Suppose we have the following linear model:

$$y_t = x_t'\beta + \epsilon_t \tag{5.11}$$

where β is a vector of unknown parameters and x_t is a vector of variables.

Suppose we have observations of size T, $x_t, x_{t+1} \cdots, x_{t+T-1}$ and $y_t, y_{t+1} \cdots, y_{t+T-1}$. Let

$$y = \begin{pmatrix} y_t \\ y_{t+1} \\ \vdots \\ y_{t+T-1} \end{pmatrix}, \quad A = \begin{pmatrix} x'_t \\ x'_{t+1} \\ \vdots \\ x'_{t+T-1} \end{pmatrix}, \quad \text{and} \quad \epsilon = \begin{pmatrix} \epsilon_t \\ \epsilon_{t+1} \\ \vdots \\ \epsilon_{t+T-1} \end{pmatrix} \tag{5.12}$$

then we have a linear system

$$\epsilon = y - A\beta \tag{5.13}$$

We need to find an approximation $\hat{\beta}$ such that ϵ is minimized. This is true if and only if the vector ϵ is orthogonal to the range of A, this implies that

$$A'\epsilon = 0 \tag{5.14}$$

This is equivalently

$$A'A\beta = A'y \tag{5.15}$$

then the OLE estimator $\hat{\beta}_{OLS}$ is given by

$$\hat{\beta}_{OLS} = (A'A)^{-1}A'y = \left(\sum_{t=1}^{T} x_t x'_t\right)^{-1} \left(\sum_{t=1}^{T} x_t y_t\right) \tag{5.16}$$

Now let's redo the estimation in the GMM framework. The critical assumption to get $\hat{\beta}_{OLS}$ is (5.14), then we can make the moment condition as

$$E(x_t \epsilon_t) = 0 \tag{5.17}$$

which is the same as the true value β_0 must satisfy

$$E[x_t(y_t - x_t'\beta_0)] = 0 \tag{5.18}$$

Let $w_t = (x_t', y_t)$, and $\theta = \beta$, we can define $h(\cdot, \cdot)$ as

$$h(\theta, w_t) = x_t(y_t - x_t'\theta) \tag{5.19}$$

note here that the number of orthogonality conditions is the same as the number of unknown parameters, so this is an exactly identified system. Thus the GMM estimator $\hat{\theta}_{GMM}$ can be found by solving

$$0 = g(\theta, Y_T) = \frac{1}{T}\sum_{t=1}^{T} x_t(y_t - x_t'\hat{\theta}_{GMM}) \tag{5.20}$$

which gives us

$$\hat{\theta}_{GMM} = \left(\sum_{t=1}^{T} x_t x_t'\right)^{-1}\left(\sum_{t=1}^{T} x_t y_t\right) \tag{5.21}$$

which is the same as the OLS estimator.

5.1.3 Maximum Likelihood Estimation. Maximum likelihood estimation is another method of estimating parameters. This method can be applied when the distribution is known. Suppose $f(x|\beta)$ is a density function of some distribution with parameters $\beta = \{\beta_1, \ldots, \beta_m\}$, and $\{x_1, \cdots, x_T\}$ is sample data from observation with that distribution. The samples are assumed to be i.i.d., then the joint density function will be

$$f(x_1, \cdots, x_T|\beta) = \prod_{i=1}^{T} f(x_i|\beta) := \mathcal{L}(\beta|x_1, \cdots, x_T) \tag{5.22}$$

For each set of parameters β, we will get a density function, which gives us the probability that the sample event happens, or in other words, the likelihood of the sample events. We call the function $\mathcal{L}$ the likelihood function, and we want to find the set of parameters that maximized the likelihood function.

Here f is a density function so β are parameters for f. On the other hand, the likelihood function will give a set of parameters with each observation, therefore the $x_1, \cdots, x_n$ are parameters for $\mathcal{L}$. In practice it is often convenient to work with the logarithm of $\mathcal{L}$ and scale it, we call it the average log-likelihood:

$$l(\beta|x_1, \cdots, x_T) = \frac{1}{T} \ln \mathcal{L}(\beta|x_1, \cdots, x_T) = \frac{1}{n} \sum_{i=1}^{T} \ln f(\beta)(x_i) \tag{5.23}$$

we need to find the vector β to maximize $l(\beta|x_1, \cdots, x_n)$, i.e.

$$\hat{\beta}_{mle} = \max_{\beta} l(\beta|x_1, \cdots, x_T) = \max_{\beta} \frac{1}{T} \sum_{i=1}^{n} \ln f(\beta)(x_i). \tag{5.24}$$

Thus we can get the MLS estimator $\hat{\theta}_{MLE}$ by solving the following equation:

$$\sum_{i=1}^{N} \frac{\partial \ln f(\beta|x_i)}{\partial \beta} = 0 \tag{5.25}$$

In the GMM framework, let $Y_T = \{x_1, \cdots, x_T\}$. Suppose that the conditional density of the tth observation is given as

$$f(y_t|Y_{t-1}; \theta) \tag{5.26}$$

since f is a density function we should have

$$\int_{\mathcal{D}} f(y_t|Y_{t-1}; \theta) dy_t = 1 \tag{5.27}$$

where $\mathcal{D}$ denotes the set of all possible values of y_t. If we differentiate (5.27) with respect to θ, we should have

$$\int_{\mathcal{D}} \frac{\partial f(y_t|Y_{t-1}; \theta)}{\partial \theta} dy_t = 0 \tag{5.28}$$

We can multiply and divide the integrand by the density f to get

$$\int_{\mathcal{D}} \frac{\partial f(y_t|Y_{t-1};\theta)}{\partial\theta} \frac{1}{f(y_t|Y_{t-1};\theta)} f(y_t|Y_{t-1};\theta)dy_t = 0 \tag{5.29}$$

which can be written as

$$\int_{\mathcal{D}} \frac{\partial \ln f(y_t|Y_{t-1};\theta)}{\partial\theta} f(y_t|Y_{t-1};\theta)dy_t = 0 \tag{5.30}$$

Let $h(\theta, Y_t)$ denote the log of the density function

$$h(\theta, Y_t) = \frac{\partial \ln f(y_t|Y_{t-1};\theta)}{\partial\theta} \tag{5.31}$$

now if we combine (5.30) and (5.31), we will get

$$E[h(\theta, Y_t)|Y_{t-1}] = 0 \tag{5.32}$$

then with the iteration condition of conditional expectation we should have

$$E[h(\theta, Y_t)] = 0 \tag{5.33}$$

we get the orthogonality condition. The GMM framework suggests using the estimator $\hat{\theta}_{GMM}$, that is obtained by solving

$$\frac{1}{T}\sum_{t=1}^{T} h(\theta, Y_t) \tag{5.34}$$

which is exactly the same conditions as (5.25). Thus MLE is the same as the GMM estimator based on the orthogonality conditions.

5.1.4 GMM estimation of the Vasicek Model.

In this dissertation, we will use the GMM to estimate the parameters in the Vasicek Model by following the approach in [23] and using a Euler form of the model

$$r_{t+1} - r_t = \lambda\theta - \lambda r_t \delta t + \epsilon_{t+1} \tag{5.35}$$

where $\epsilon_{t+1} = \sigma(W_{t+1} - W_t)$ is a normally distributed random variable with mean zero and variance $\sigma^2 \Delta t$. Then we have the first two moment conditions

$$E[\epsilon_{t+1}] = 0 \quad \text{and} \quad E[\epsilon_{t+1}^2 - \sigma^2 \Delta t] = 0 \tag{5.36}$$

Note that the value of r_t is independent of ϵ_{t+1}, thus we should have two more moment conditions

$$E[\epsilon_{t+1} r_t] = 0 \quad \text{and} \quad E[(\epsilon_{t+1}^2 - \sigma^2 \Delta t) r_t] = 0 \tag{5.37}$$

these four conditions are used in [23]. Since W_t is Brownian motion, thus we know the increments are also independent, then I will add one more moment condition

$$E[\epsilon_{t+1} \epsilon_t] = 0 \tag{5.38}$$

then I will define $\beta = (\lambda, \theta, \sigma)$ to be the vector of the parameters, $w_t = (r_{t-1}, r_t, r_{t+1})$ to be the vector of the observations and the function h to be

$$h(\theta, w_t) = \begin{bmatrix} \epsilon_{t+1} \\ \epsilon_{t+1}^2 - \sigma^2 \Delta t \\ \epsilon_{t+1} r_t \\ (\epsilon_{t+1}^2 - \sigma^2 \Delta t) r_t \\ \epsilon_{t+1} \epsilon_t \end{bmatrix} \tag{5.39}$$

5.1.5 MLE estimation of the Vasicek Model. The solution for the Vasicek Model is given as

$$r(t) = r(0)e^{-\theta t} + \lambda \left(1 - e^{-\theta t}\right) + \sigma \int_0^t e^{-\theta(t-u)} dW_u \tag{5.40}$$

we can calculate the distribution of the term $\sigma\int_0^t e^{-\theta(t-u)}dW_u$ to get it is normally distributed with mean zero and variance

$$\hat{\sigma}^2 = \frac{\sigma^2}{2\lambda}(1 - e^{-\lambda t}) \tag{5.41}$$

thus we can rewrite the solution as

$$r(t) = r(0)e^{-\theta t} + \lambda\left(1 - e^{-\theta t}\right) + \hat{\sigma}\epsilon \tag{5.42}$$

where $\epsilon \sim \mathcal{N}(0,1)$.

Let $r(t_0), r(t_1), \ldots, r(t_n)$ be an observation, with time step Δt. Let r_i denote $r(t_i)$, then we have

$$r_{i+1} = r_i e^{-\theta\Delta t} + \lambda\left(1 - e^{-\theta\Delta t}\right) + \hat{\sigma}\epsilon \tag{5.43}$$

and

$$\hat{\sigma}^2 = \frac{\sigma^2}{2\lambda}(1 - e^{-\lambda\Delta t}) \tag{5.44}$$

then the conditional density function of r_{i+1} given r_i is

$$f(r_{i+1}|r_i;\lambda,\theta,\hat{\sigma}) = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}}\exp\left\{-\frac{(r_{i+1} - r_i e^{-\lambda\Delta t} - \theta(1 - e^{-\lambda\Delta t}))^2}{2\hat{\sigma}^2}\right\} \tag{5.45}$$

then the log-likelihood function is

$$g(r_{i+1}|r_i;\lambda,\theta,\hat{\sigma}) = -\frac{n}{2}\log 2\pi - n\log\hat{\sigma} - \frac{1}{2\hat{\sigma}}\sum_{i=0}^{n-1}\left(r_{i+1} - r_i e^{-\lambda\Delta t} - \theta(1 - e^{-\lambda\Delta t})\right)^2 \tag{5.46}$$

To maximize $g(r_{i+1}|r_i;\lambda,\theta,\hat{\sigma})$, we take its first partial derivatives with respect to $\lambda, \theta, \hat{\sigma}$, set them equal to zero, then we have the following estimation

$$\lambda = -\frac{1}{\Delta t}\log\left\{\frac{\sum_{i=1}^n r_i^2 - \theta\sum_{i=0}^{n-1} r_i - \theta\sum_{i=1}^n r_i + n\theta^2}{\sum_{i=0}^{n-1} r_i^2 - 2\theta\sum_{i=0}^{n-1} r_i + n\theta^2}\right\} \tag{5.47}$$

$$\theta = \frac{\sum_{i=1}^n r_i\sum_{i=0}^{n-1} r_i^2 - \sum_{i=0}^{n-1} r_i\sum_{i=1}^n r_i^2}{n(\sum_{i=0}^{n-1} r_i^2 - \sum_{i=1}^n r_i^2) - (\sum_{i=0}^{n-1} r_i^2 - \sum_{i=0}^{n-1} r_i\sum_{i=1}^n r_i)} \tag{5.48}$$

$$\hat{\sigma}^2 = \frac{\sum_{i=0}^{n-1}(r_{i+1} - r_i e^{-\lambda\Delta t} - \theta(1 - e^{-\lambda\Delta t}))^2}{n} \tag{5.49}$$

5.2 Numerical Computation

In this section, we will use Monte Carlo simulation methods to price a spread option with stochastic interest rate model. The interest rate is assumed to follow the Vasicek model.

Recall the SDE for the price of two underlying assets are

$$\begin{aligned} \frac{dS_{1,t}}{S_{1,t}} &= r(t)dt + \sigma_1 d\widetilde{W}_{1,t} \\ \frac{dS_{2,t}}{S_{2,t}} &= r(t)dt + \sigma_2 d\widetilde{W}_{2,t} \end{aligned} \tag{5.50}$$

and the solutions are given as

$$\begin{aligned} S_{1,t} &= S_{1,0} \exp\left(\int_0^t r(u)du - \frac{1}{2}\sigma_1^2 t + \sigma_1 W_{1,t}\right) \\ S_{2,t} &= S_{2,0} \exp\left(\int_0^t r(u)du - \frac{1}{2}\sigma_2^2 t + \sigma_2 W_{2,t}\right) \end{aligned} \tag{5.51}$$

where $r(t)$ satisfies

$$dr(t) = \lambda(\theta - r(t))dt + \sigma_3 dW_{3,t} \tag{5.52}$$

First, we need to simulate the integral of the interest rate $\int_0^t r(u)du$. We use the Euler scheme in [24] to discretize the SDE in time for $r(t)$. Let Δt denote the step size of the discretization in time, so we have the sample points $0 = t_1 < t_2 < \cdots < t_n$, then for $j = 1, 2, \ldots, n-1$, the scheme is given as

$$r_{t_{j+1}} = r_{t_j} + \lambda(\theta - r_{t+j})\delta t + \sigma_3 r_{t+j}^{\gamma} \Delta W_{3,j+1} \tag{5.53}$$

where

$$\Delta W_{3,j+1} = W_{3,t_{j+1}} - W_{3,t_j} \tag{5.54}$$

then use the trapezoid rule to calculate the integral. Figure (5.1) The convergence of the Monte Carlo methods for the Vasicek Model.

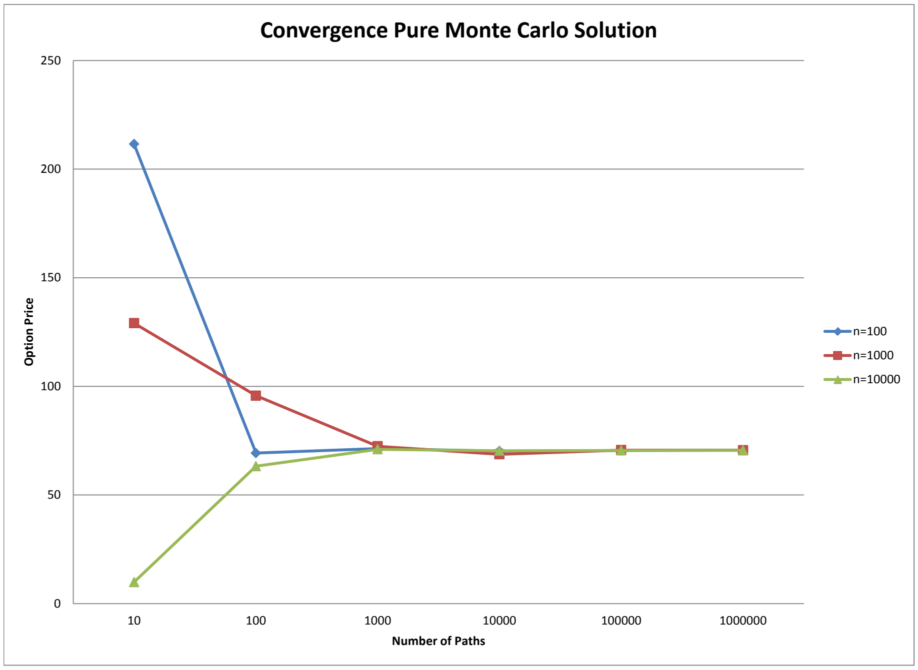

Figure 5.1: Convergence of Pure Monte Carlo, Vasicek Model

Table (5.2) shows the full result of the simulation for Vasicek Model.

	N=100		N=1000		N=10000	
M	Price	Time	Price	Time	Price	Time
10	211.498806	0.457882	129.094167	0.003038	9.887732	0.028902
100	69.304983	0.005625	95.719170	0.031521	63.269905	0.290680
1000	71.376245	0.057146	72.427088	0.286408	71.032658	2.836017
10000	70.219813	0.488674	68.735862	2.871948	70.369001	27.575780
100000	70.472602	5.042892	70.645601	27.937238	70.495426	275.407357
1000000	70.617125	48.046817	70.583438	280.210989	70.560143	2753.550529

Table 5.1: Computation time for Pure Monte Carlo Solution, Vasicek Model

As you can see in the table (5.2) and figure (5.1) , to get the convergence, we need more than 50000 paths, this is because by the Central Limit Theorem the error of Monte Carlo simulation with m sample paths has standard deviation $\frac{\sigma}{\sqrt{m}}$, where σ is the true standard divination, which is usually approximated by the sample standard deviation σ_n. In order to improve the error by 0.1 we need 100 more sample paths. This is really costly on computational time. So in order to reduce the number of sample paths and the computational time, we can reduce the σ_n, this is called variance reduction. There are many variance reduction methods, in this dissertation, we will use the Anti-thetic method in [25]. The idea of anti-thetic is the following:

Suppose X is a random variable, and let $Y = h(X)$, and we want to estimate $\theta = E[Y]$. Let $\{x_1, \cdots, x_m\}$ be a sample path, the pure Monte Carlo method estimator is given by

$$\hat{\theta}_{pmc} = \frac{1}{m}\sum_{i=1}^{m} h(x_i) \tag{5.55}$$

for the anti-thetic method, we also use the antithetic path $\{-x_1, \cdots, -x_m\}$. Let

$$\hat{\theta}_1 = \frac{1}{m}\sum_{i=1}^{m} h(x_i) \tag{5.56}$$

and

$$\hat{\theta}_2 = \frac{1}{m} \sum_{i=1}^{m} h(-x_i) \tag{5.57}$$

then the anti-thetic estimation for θ is given as

$$\hat{\theta}_{anti} = \frac{\hat{\theta}_1 + \hat{\theta}_2}{2} \tag{5.58}$$

the variance of this estimator is given

$$\begin{aligned} Var(\hat{\theta}_{anti}) &= \frac{Var(\hat{\theta}_1) + Var(\hat{\theta}_2) + 2Cov(\hat{\theta}_1, \hat{\theta}_2)}{4} \\ &= \frac{2Var(\hat{\theta}_1) + 2Cov(\hat{\theta}_1, \hat{\theta}_2)}{4} \end{aligned} \tag{5.59}$$

here we use the fact that $\hat{\theta}_1$ and $\hat{\theta}_2$ have the same distribution, since we just use the thetic path. Then the variance is reduced due to the fact that the covariance is less than or equal to the variance. Figure (5.2) and table (5.2) show the simulation results for the anti-thetic method:

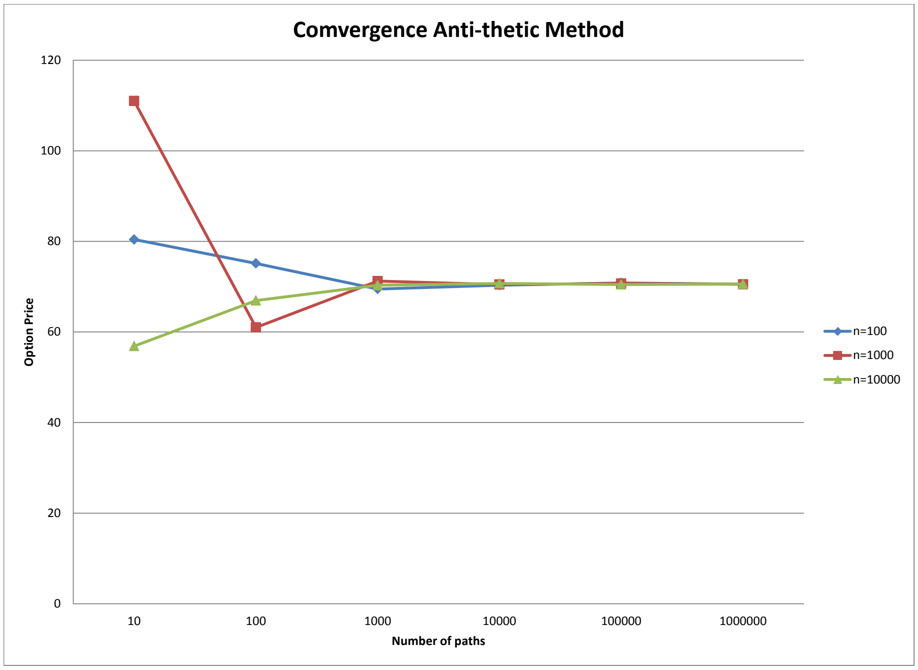

Figure 5.2: Convergence of Monte Carlo Anti-thetic Method, Vasicek Model

We also used two Taylor-Expansion schemes, Euler scheme and Milstein scheme to simulate the value of $S_{1,T}$ and $S_{2,T}$ with the anti-thetic method. the Results are shown table (5.2), (5.2) and figure (5.3), (5.4).

	N=10		N=100		N=1000	
M	Price	Time	Price	Time	Price	Time
10	80.418102	0.001467	111.014902	0.012199	56.868863	0.036838
100	75.142338	0.006729	61.013739	0.043828	66.945374	0.416838
1000	69.475202	0.068035	71.224411	0.388719	70.306222	3.686895
10000	70.323580	0.760436	70.467464	3.739498	70.722988	35.146937
100000	70.832241	6.467662	70.698599	36.199116	70.465143	356.198136
1000000	70.538551	60.001175	70.526864	357.789468	70.576108	3585.312342

Table 5.2: Computation time for Monte Carlo Anti-thetic Method, Vasicek Model

	N=10		N=100		N=1000	
M	Price	Time	Price	Time	Price	Time
10	79.020419	0.001524	97.507355	0.005457	46.604596	0.047006
100	63.889017	0.008355	65.055875	0.044977	67.157119	0.452856
1000	72.284993	0.073500	71.579314	0.472439	71.074498	4.262531
10000	69.045844	0.720953	72.444331	4.288694	71.317078	43.532677
100000	70.504883	6.911399	70.400884	43.431433	70.225730	438.838317
1000000	70.384021	68.836795	70.606029	430.485230	70.398284	4302.645862

Table 5.3: Computation time for Monte Carlo Anti-thetic Method Euler, Vasicek Model

We can see the clear picture in table 5.2. With $n = 1000$, for each fixed m, we ran the simulation twice and recorded the difference between the two prices. Table 5.2 shows the price difference and time needed for each simulation. We can get a better sense that the Monte Carlo simulation converged very slowly.

	N=10		N=100		N=1000	
M	Price	Time	Price	Time	Price	Time
10	39.260039	0.001708	67.846614	0.007915	63.552036	0.060463
100	63.166351	0.009483	75.808585	0.056915	60.311508	0.579815
1000	68.630529	0.087046	68.280957	0.566781	70.432008	5.469940
10000	70.017462	0.865116	69.279721	5.532248	70.616620	55.433094
100000	70.301982	8.533419	70.131954	55.981405	70.776102	558.264598
1000000	70.248217	83.993073	70.351437	559.934416	70.506756	5536.634632

Table 5.4: Computation time for Monte Carlo Anti-thetic Method Milstein, Vasicek Model

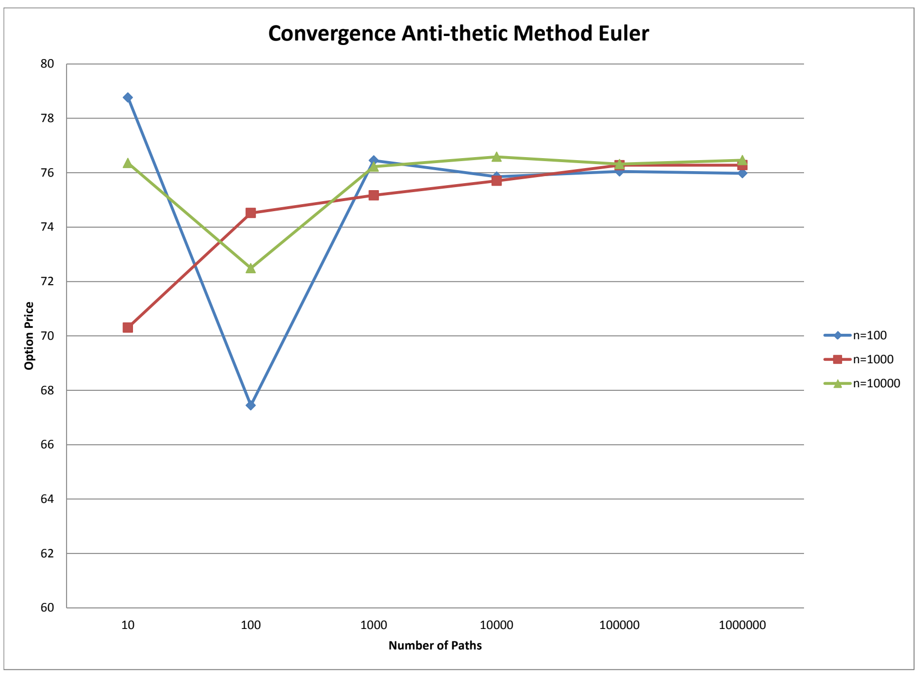

Figure 5.3: Convergence of Monte Carlo Anti-thetic Euler, Vasicek Model

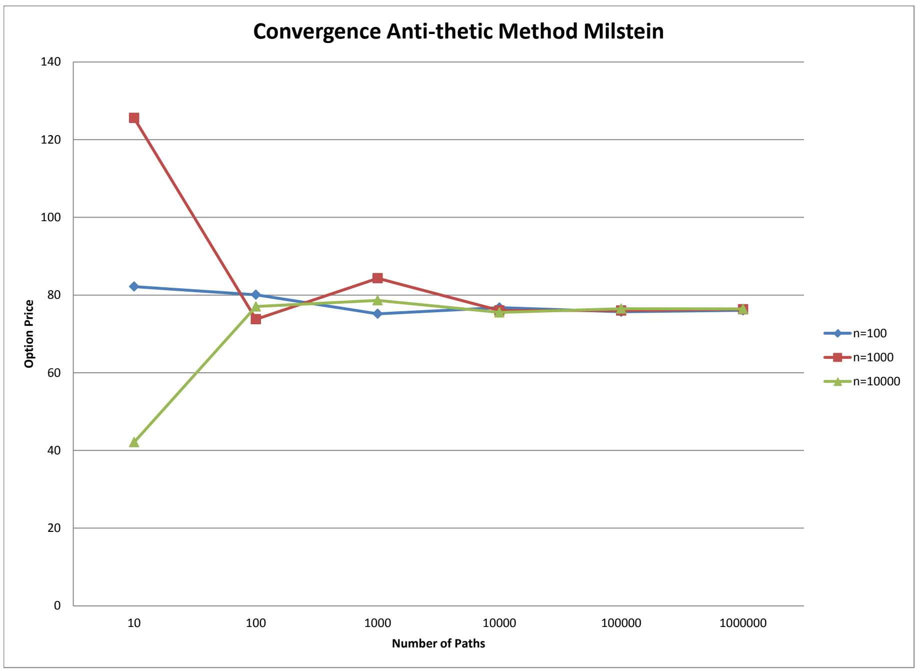

Figure 5.4: Convergence of Monte Carlo Anti-thetic Milstein, Vasicek Model

M	Error	Time (Seconds)
10000	3.067547	3.11
50000	0.676	14.46
250000	0.466846	70.67
1250000	0.050257	352.03
6250000	0.047147	1757.59
31250000	0.027583	8769.58

Table 5.5: Errors for Pure Monte Carlo Simulation with $n = 1000$

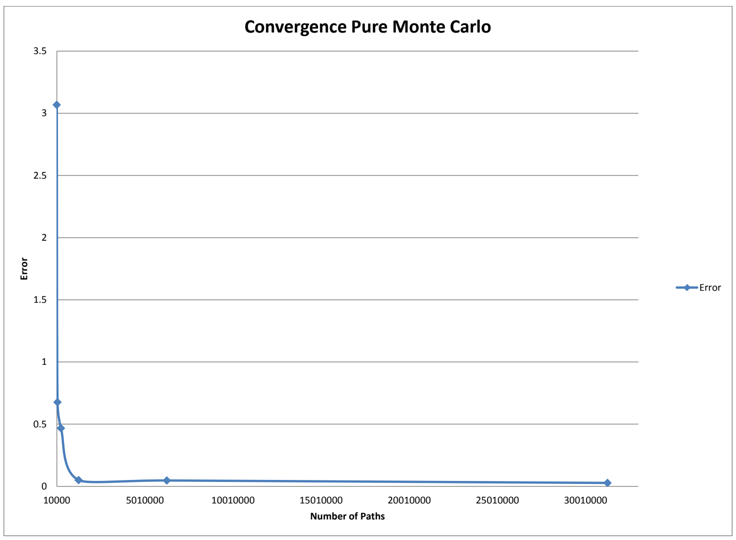

Figure 5.5: Errors for the Pure Monte Carlo

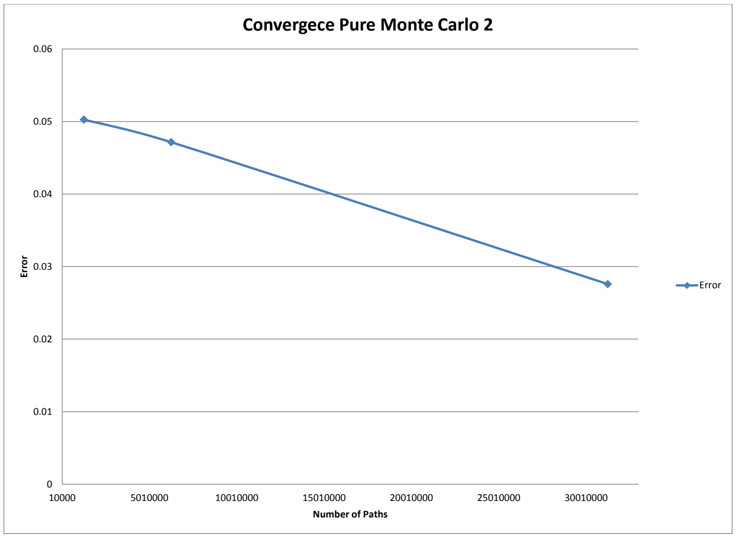

Figure 5.6: Errors for the Pure Monte Carlo

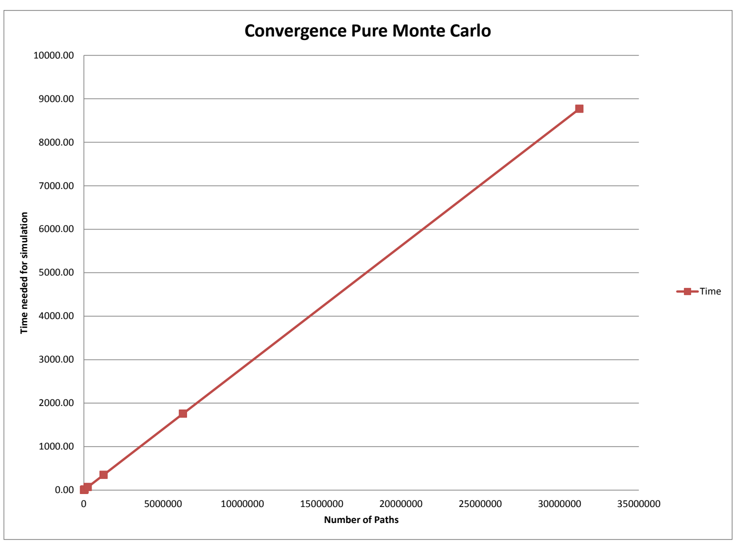

Figure 5.7: Time Required for the Pure Monte Carlo

Table 5.2 shows that with anti-thetic method, with same number of sample paths, the price converged faster. Figures (5.10) and (5.11) shows the comparison of the error for the two methods.

M	Error	Time (Seconds)
10000	0.835195	4.090361
50000	0.188076	21.093107
250000	0.076690	93.179374
1250000	0.047806	454.945275
6250000	0.006680	2231.995729

Table 5.6: Errors for Monte Carlo Simulation Anti-thetic method with $n = 1000$

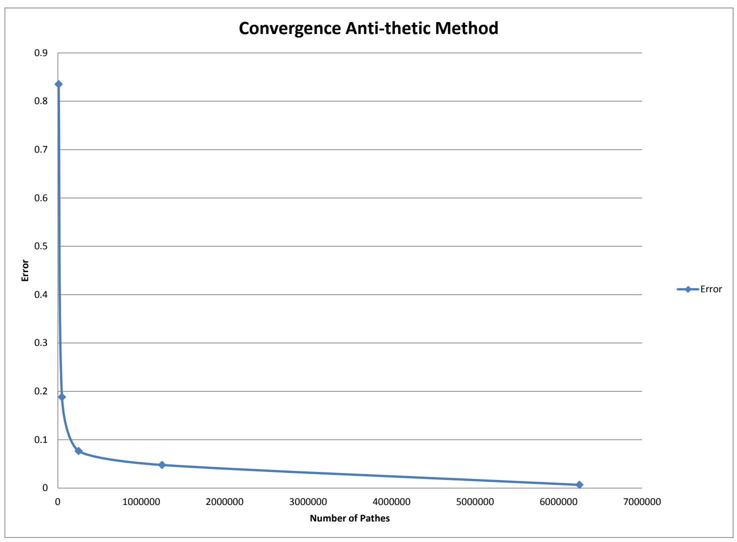

Figure 5.8: Errors for the Pure Monte Carlo Anti-thetic method

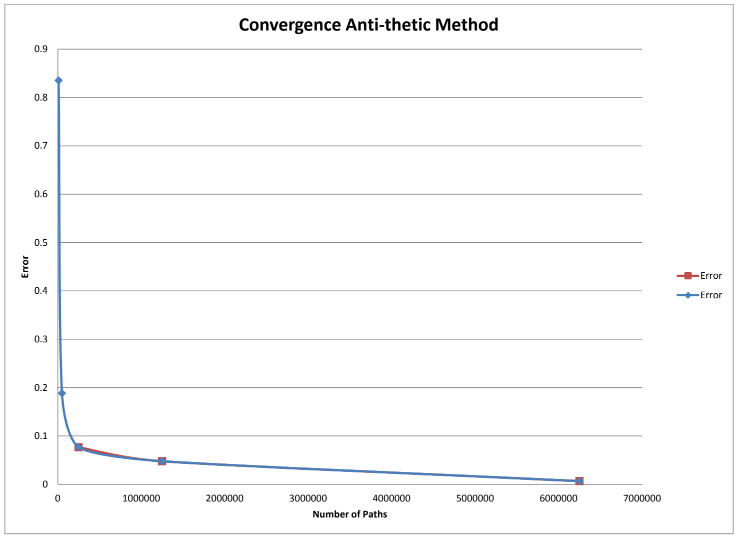

Figure 5.9: Errors for the Monte Carlo Anti-thetic method

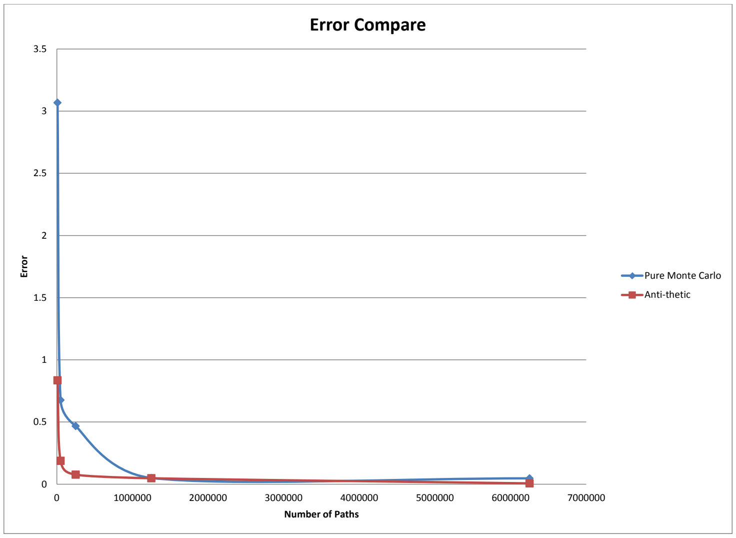

Figure 5.10: Error Comparison of PMC and Anti-thetic Methods

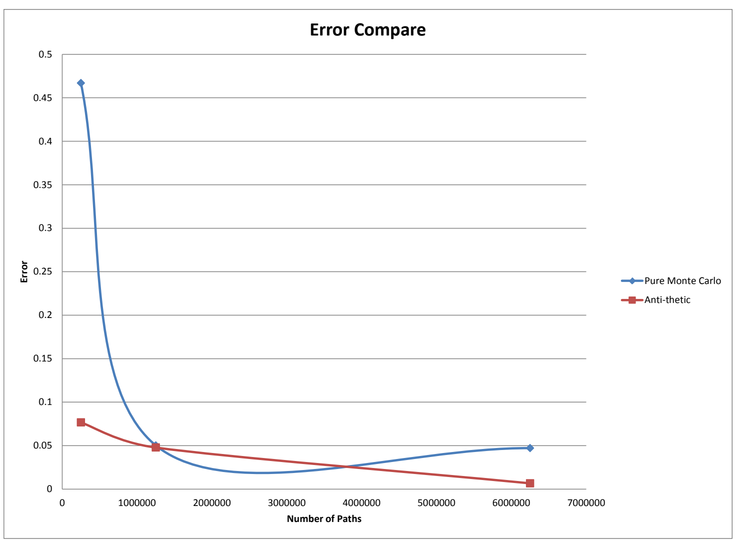

Figure 5.11: Error Comparison of PMC and Anti-thetic Methods

Note in figure (5.12) and (5.12) that the time required for the simulation for the anti-thetic methods is longer than the pure Monte Carlo methods, this is due to the fact that we have to perform the calculation twice. But compared to the significance of the error being reduced, the increasing of the simulation time can be ignored.

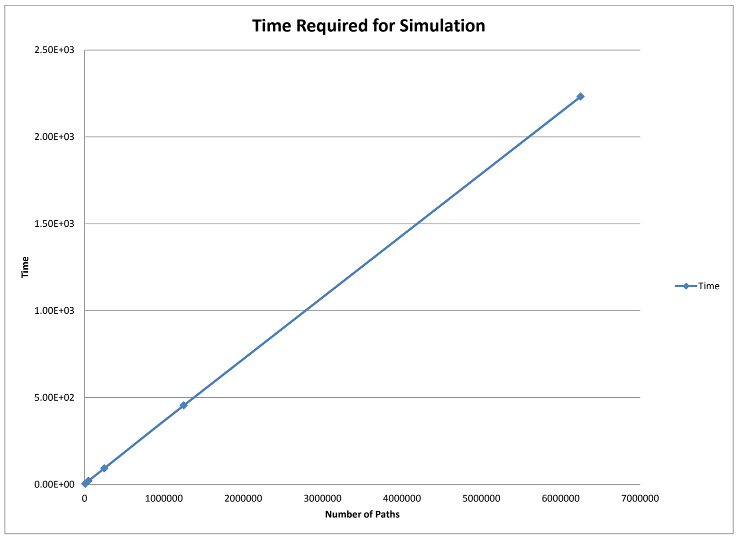

Figure 5.12: Time Required for the Monte Carlo Anti-thetic Method

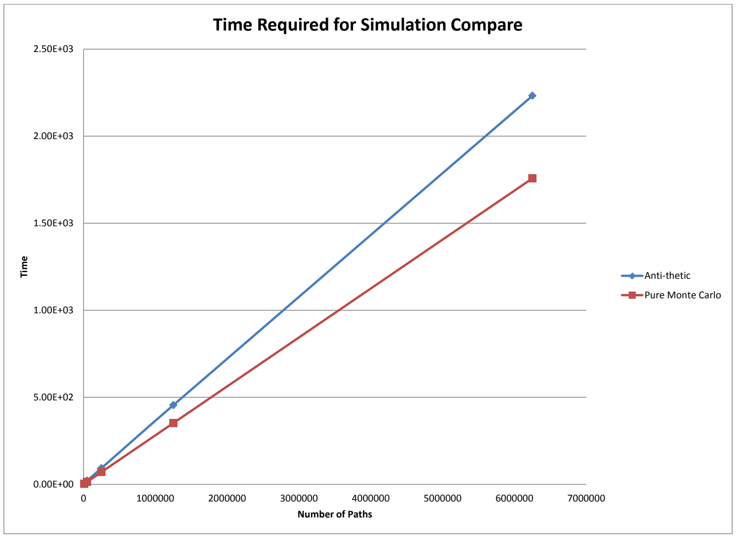

Figure 5.13: Comparison of Time Required for the PMC and Anti-thetic Methods

λ	Constant Interest Rate	Anti-thetic	Euler	Milstein
0.000000	70.547906	110.464054	110.608279	110.209345
0.600000	70.716412	106.218147	106.025431	106.044925
5.000000	70.702931	92.553106	92.451512	92.394744
10.000000	70.903979	88.165538	88.350379	88.218140
50.000000	70.576363	84.067635	84.053208	84.191751

Table 5.7: Price Compare with Different Values of λ with $r_0 = 0.0008,\ \theta = 0.0534$

One of the natural questions to ask is that under the same parameters, how does the price in the stochastic interest rate model compare to the constant rate model; the following tables shows the result. I compared the prices of constant interest rate and stochastic interest rate, under different market environment by changing the parameters. Table (5.2) shows the result for different values of the mean reverting speed λ when the initial interest rate is less than θ, table (5.2) shows the result when initial interest rate is larger than θ. Table (5.2) shows the result for different values of θ, and Table (5.2) shows the result for the different values of the volatility of interest rate.

λ	Constant Interest Rate	Anti-thetic	Euler	Milstein
0.000000	94.515013	129.811666	129.775950	129.663419
0.600000	94.349741	121.442713	120.994013	121.521613
5.000000	94.567394	96.804342	96.910704	97.027675
10.000000	94.261885	90.594437	90.527466	90.409889
50.000000	94.276119	84.542643	84.390976	84.424778

Table 5.8: Price Compare with Different Values of λ with $r_0 = 0.1,\ \theta = 0.0534$

θ	Constant Interest Rate	Anti-thetic	Euler	Milstein
0.000500	70.806360	103.339074	103.479492	103.666390
0.050000	71.111960	106.255414	106.105306	105.999706
0.100000	70.870805	108.735652	108.718374	108.247126
0.150000	70.404410	110.934148	110.879887	110.803613
0.500000	70.925689	128.892505	128.750617	128.825463

Table 5.9: Price Compare with Different Values of θ

θ	Constant Interest Rate	Anti-thetic	Euler	Milstein
0.050000	70.837975	70.075135	70.008829	69.915111
0.100000	70.907698	66.332668	66.380806	66.349725
0.500000	70.819627	40.522019	40.412696	40.517000
0.900000	71.328635	35.782832	35.679400	35.753569

Table 5.10: Price Compare with Different Values of σ_3, $r_0 = 0.0008$

θ	Constant Interest Rate	Anti-thetic	Euler	Milstein
0.050000	94.359533	88.109150	88.267598	88.035069
0.100000	94.447854	84.703259	84.520928	84.811073
0.500000	94.483463	59.672619	59.633542	59.641797
0.900000	94.397265	51.354077	51.383000	51.219548

Table 5.11: Price Compare with Different Values of σ_3, $r_0 = 0.1$

λ	Constant Interest Rate (0.0534)	Anti-thetic	Euler	Milstein
0.000000	83.029420	110.607722	110.695108	110.483740
0.600000	82.847328	106.235577	106.078435	106.224554
5.000000	83.009157	92.479496	92.361642	92.231534
10.000000	82.950025	88.327809	88.289298	88.174962
50.000000	83.077998	83.996396	84.068368	84.184598

Table 5.12: Price Compare with Different Values of λ with $r_0 = 0.0008,\ \theta = 0.0534$

θ	Constant Interest Rate	Anti-thetic	Euler	Milstein
1	70.624326	70.913284	71.057949	70.818340
2	70.809709	70.540870	70.337176	70.642211
3	79.753596	79.877574	79.786341	79.804307
4	88.432774	88.349079	88.541439	88.434559

Table 5.13: Price Compare with Different Values of Correlation

From table (5.2) and table (5.2) we can see the option price decreases as λ increases. Another interesting phenomenon we can see from these two tables is that, as λ increasing, the option price tends to converge to a fixed price. This is because when the reverting speed is getting larger, it would take less time for the interest rate to revert back to the mean. Once the interest rate is very close to the mean, the randomness will not drive the interest rate too far away from the mean, so the option price will converge to the price with a constant rate θ. As you can see in table (5.2), when the interest rate is set to the long term mean 5.34%, the price is around 83 dollars. When the interest rate is stochastic, the option price converges to a neighborhood of this price as λ increases.

Table (5.2) shows the price is a increasing function of the parameter θ,and Table (5.2), shows that the price is a decreasing function of the volatility of interest rate. Table (5.2), shows the difference price with different correlation between the two underlying assets and interest rate.

Table (5.2), shows the comparison for different trend of interest rate. As shown in the table, the price of an option in the stochastic interest rate model is more expensive than the constant rate when the interest rate has the trend to increasing during the life of the option,

r_0	Constant Interest Rate	Anti-thetic	Euler	Milstein
0.1	94.426384	91.732143	91.563931	91.820965
0.03	77.433120	78.293567	78.373583	78.296190

Table 5.14: Price Compare with Trend of Interest rate

and cheaper when the interest rate has the trend to decrease.

Table (5.2) and (5.2) shows the price with different strike price. We can see that the difference between the price under constant rate model and the stochastic model is small when the option is deep in the money, and it is increasing as the option is getting closer to at the money. When the option is at the money, the difference is the greatest. When the option is way out of the money, since the price is close to zero, the difference is small again. From this, we know that the stochastic interest rate model would be useful to price an at the money or close to at the money option.

K	Constant Interest Rate	Anti-thetic	Euler	Milstein
100	402.342741	399.867952	399.742870	399.782775
450	95.937442	64.031293	64.029624	64.012802
500	71.062679	37.046974	36.789771	36.804783
650	25.626759	4.174899	4.159539	4.163855
900	3.846759	0.037288	0.036218	0.035022

Table 5.15: Price Compare with Different Strike Price, $r_0 = 0.0008$

Table (5.2) and (5.2) shows the option price with change of the time to maturity. As we can see under the stochastic interest rate model, the option price is not as sensitive to the time as the constant rate model.

K	Constant Interest Rate	Anti-thetic	Euler	Milstein
100	411.861407	407.118525	407.216287	407.181275
450	122.353716	87.035146	87.041543	86.896878
500	94.307846	55.314789	55.297467	55.415321
650	39.400723	8.895968	8.835930	8.857297
900	7.592258	0.133104	0.129227	0.126691

Table 5.16: Price Compare with Different Strike Price, $r_0 = 0.1$

T	Constant Interest Rate	Anti-thetic	Euler	Milstein
0.1	23.297136	21.641312	21.627673	21.623824
0.3	39.440664	31.737008	31.606571	31.611031
0.5	50.524270	35.264094	35.252750	35.185694
0.7	59.351781	36.288908	36.292419	36.252544
0.9	67.198903	36.753626	36.742444	36.728296

Table 5.17: Price Compare with Different Time to Maturity, $r_0 = 0.0008$

T	Constant Interest Rate	Anti-thetic	Euler	Milstein
0.1	25.807930	24.143973	24.084311	24.108108
0.3	46.757322	38.636738	38.668107	38.716564
0.5	62.685391	46.340496	46.294863	46.335020
0.7	76.266826	50.921601	51.040629	50.993977
0.9	88.784190	54.082715	54.107280	54.217679

Table 5.18: Price Compare with Different Time to Maturity, $r_0 = 0.1$

Chapter 6. Conclusion and Future Work

Now we have a better understanding of spread option when the interest rate is stochastic. Under the risk-neutral model, and using the Forward measure, we have a partial differential equation for the price with free boundary condition. We can also compute the price numerically by Monte Carlo simulation. From the simulation, we know that the option price is decreasing function of the interest rate volatility, long term mean and the mean reverting speed. When the interest rate is increasing during the life of the option, the price under the stochastic model is cheaper than the price under the constant rate. On the other hand, when the interest rate is decreasing during the life of the option, constant rate model gives a cheaper price.

There are many interesting questions for the future research for this model? In this dissertation, we only used the Vasicek model for the interest rate, which allows the interest rate to be negative. How will the price change if we adapt a different term structure model such as the CIR model or Hull-White model. Another question we can ask is under what market condition, we should use a stochastic interest rate model instead of a constant rate model. We see some interesting phenomenon, what are the financial and mathematical explanation for them?

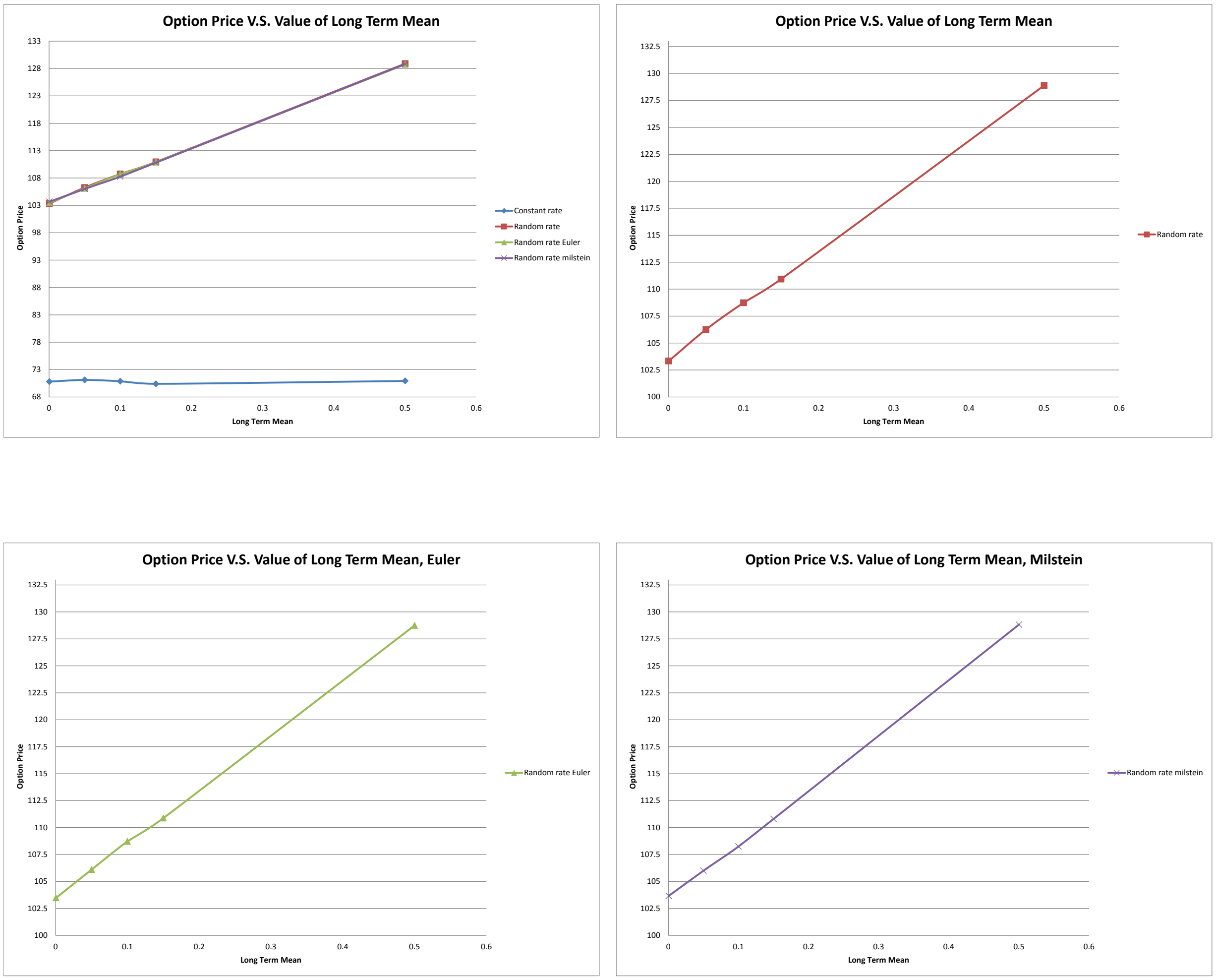

Figure 1: Option Price V.S Long Term Mean

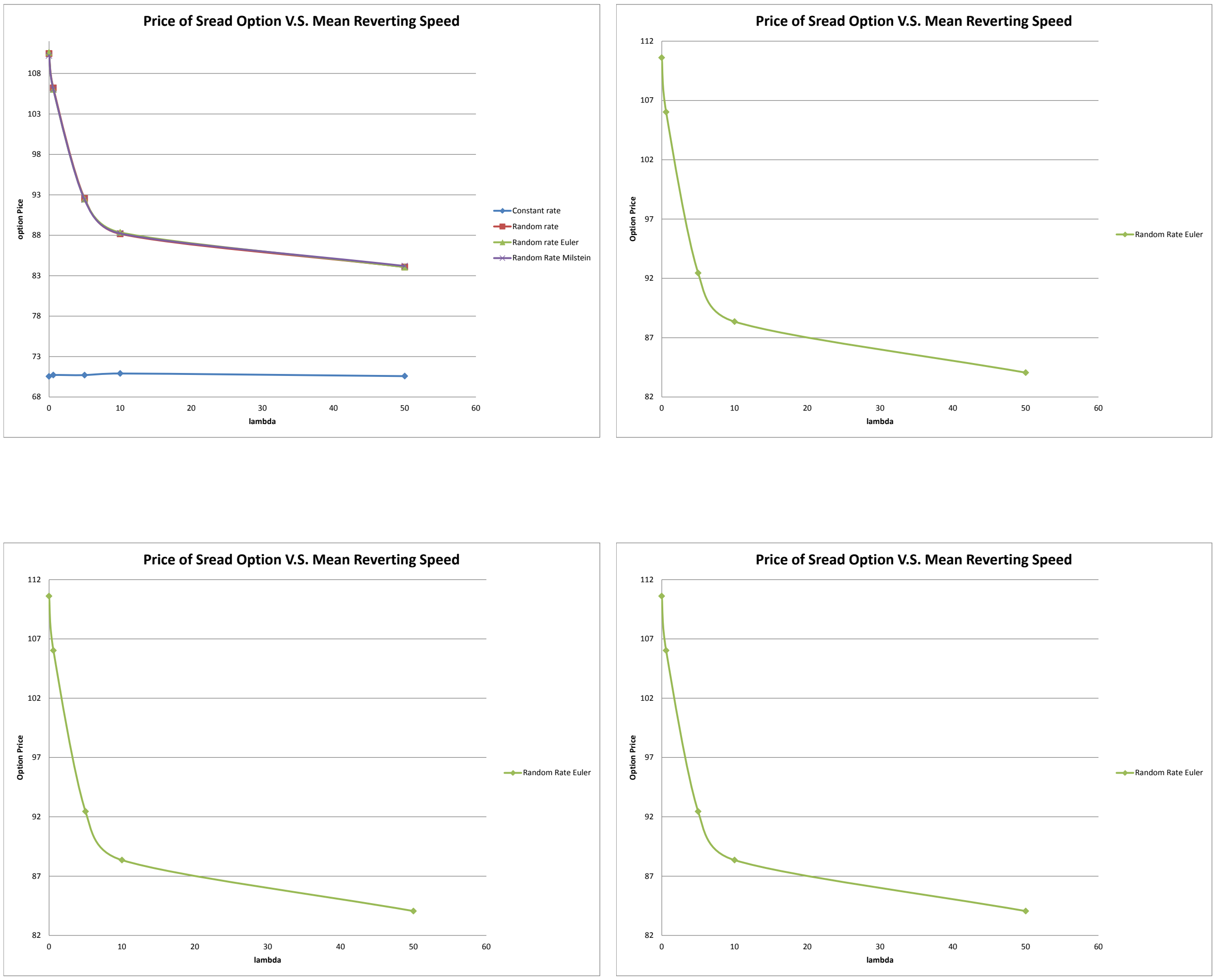

Figure 2: Option Price V.S. Mean Reverting Speed, $r_0 < \theta$

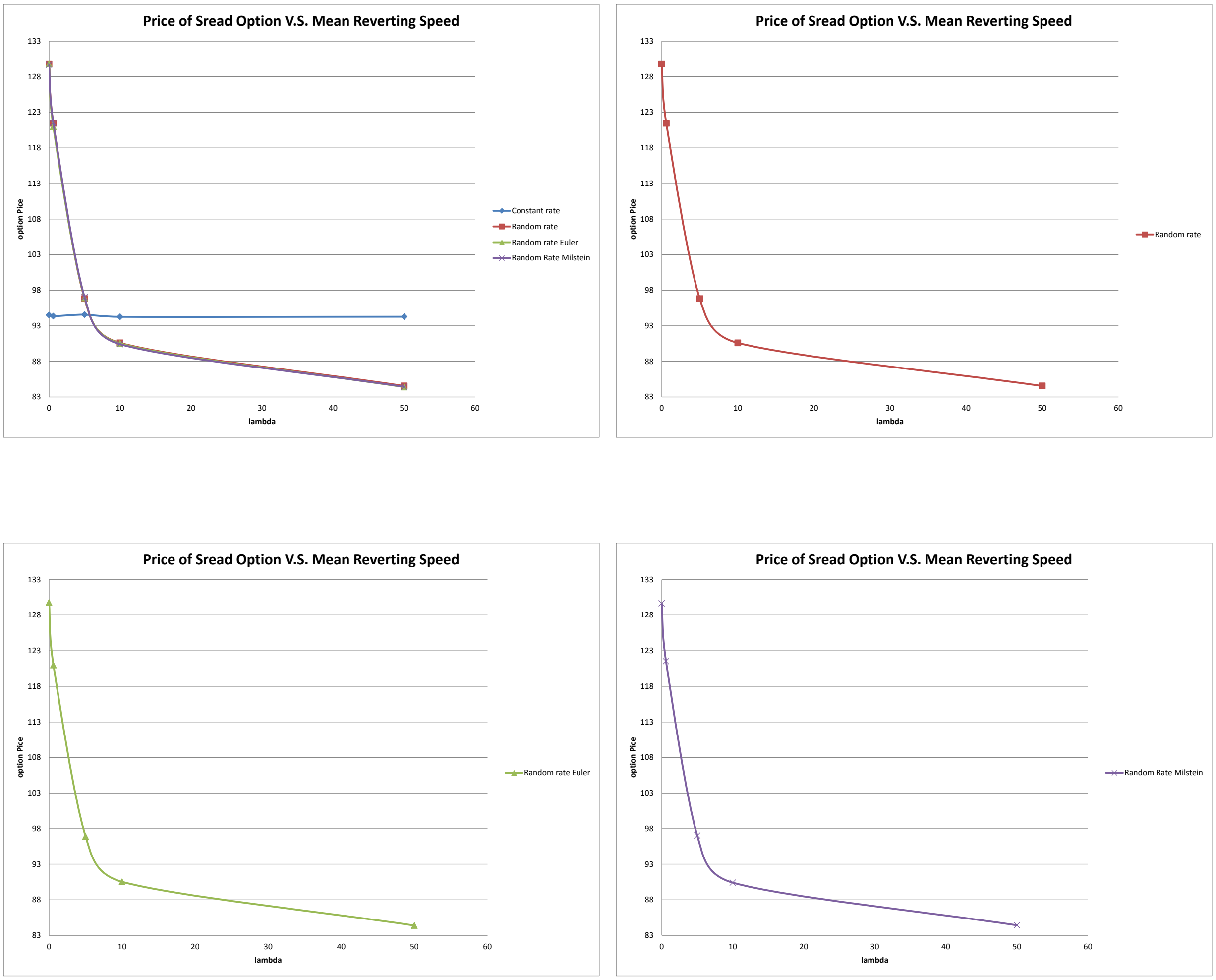

Figure 3: Option Price V.S. Mean Reverting Speed, $r_0 > \theta$

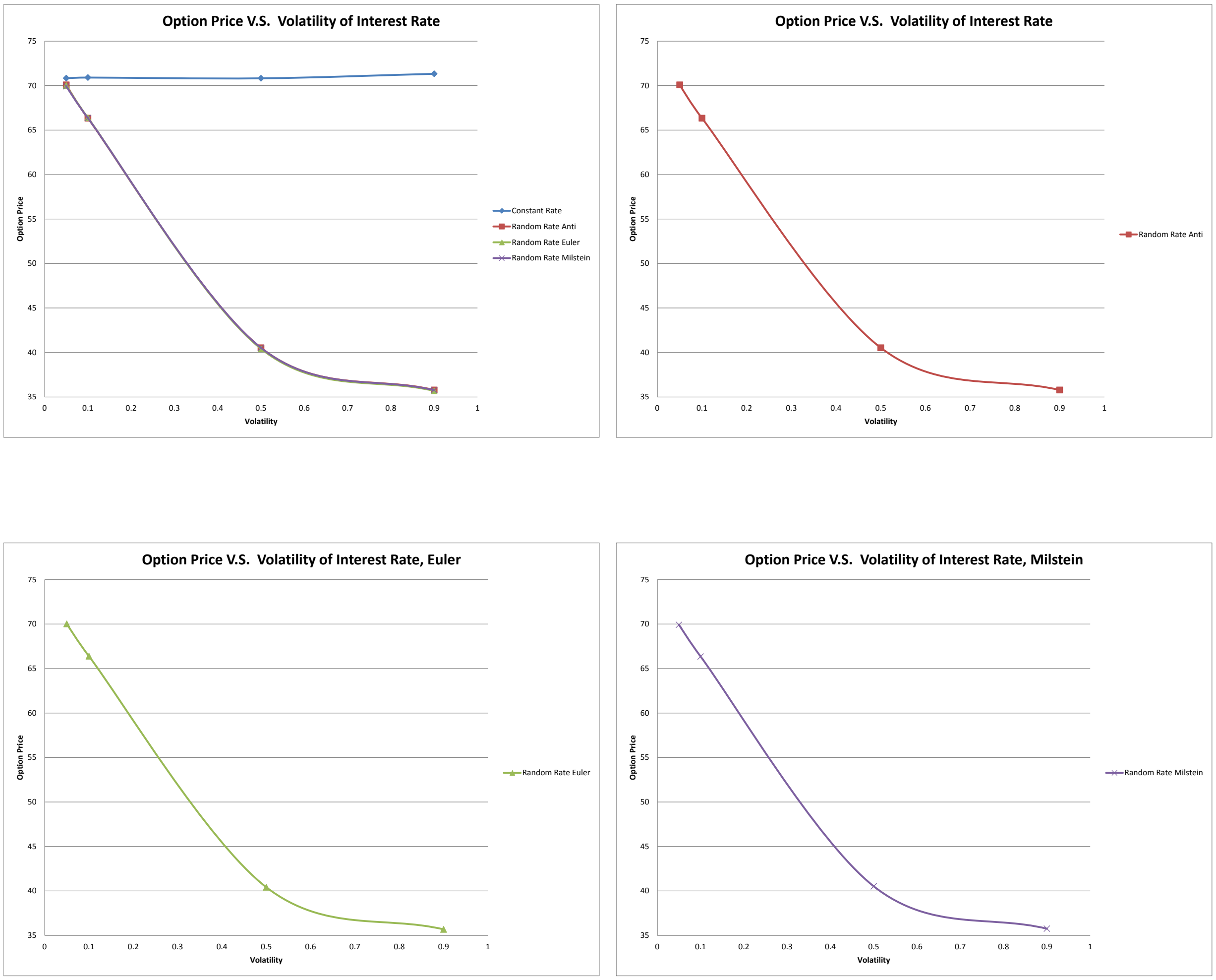

Figure 4: Option Price V.S. σ_3, $r_0 = 0.0008$

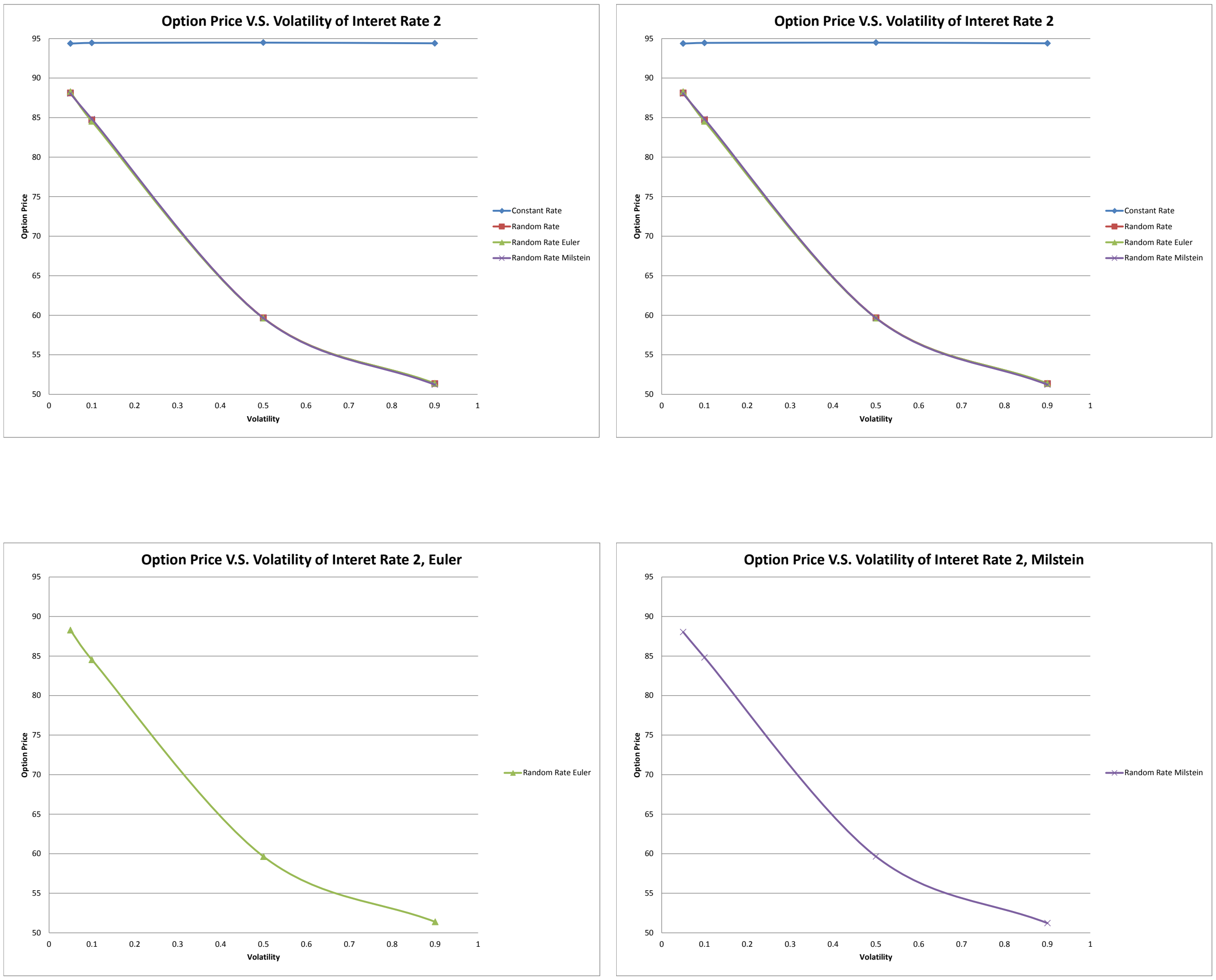

Figure 5: Option Price V.S. σ_3, $r_0 = 0.1$

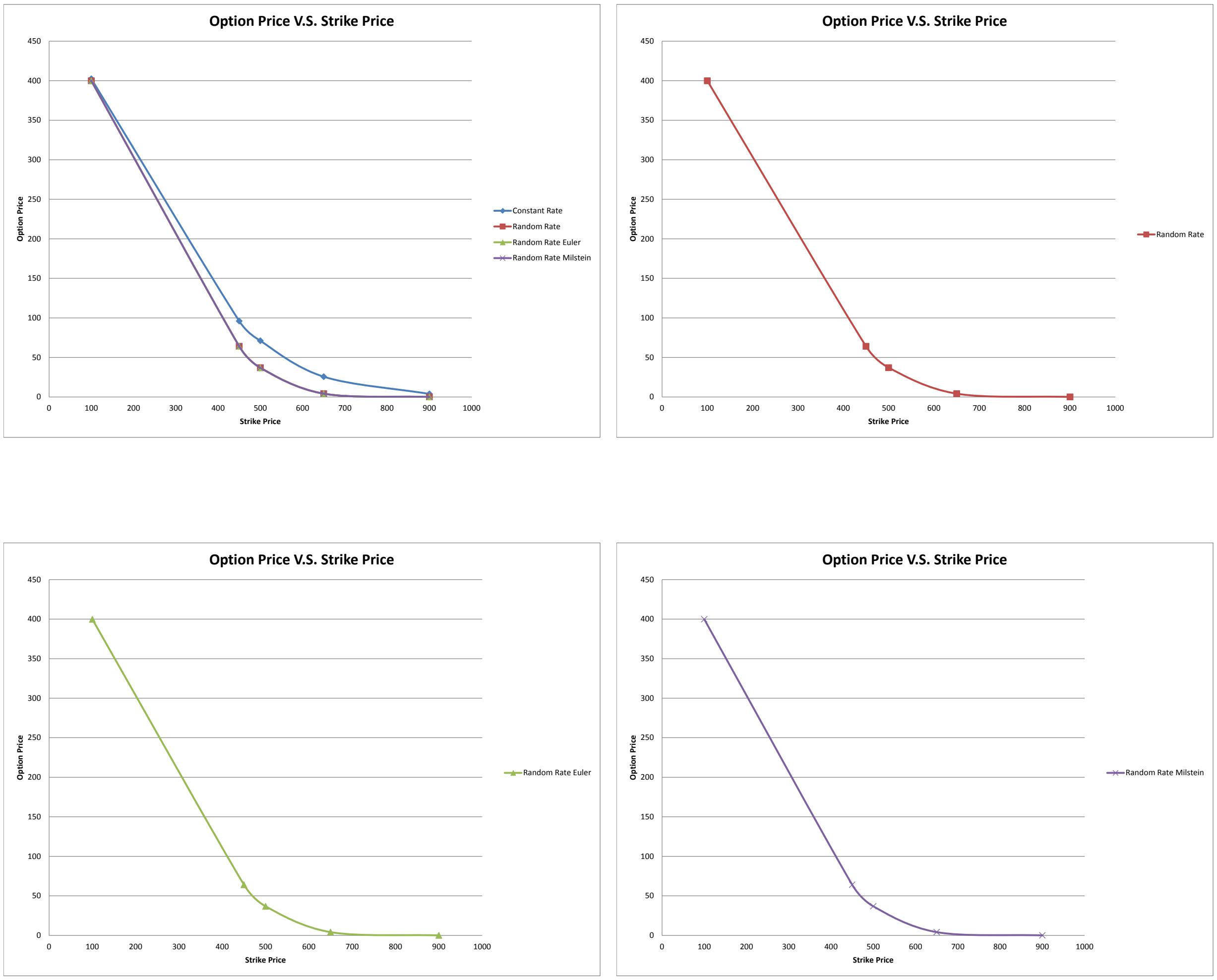

Figure 6: Option Price V.S. Strike Price, $r_0 = 0.0008$

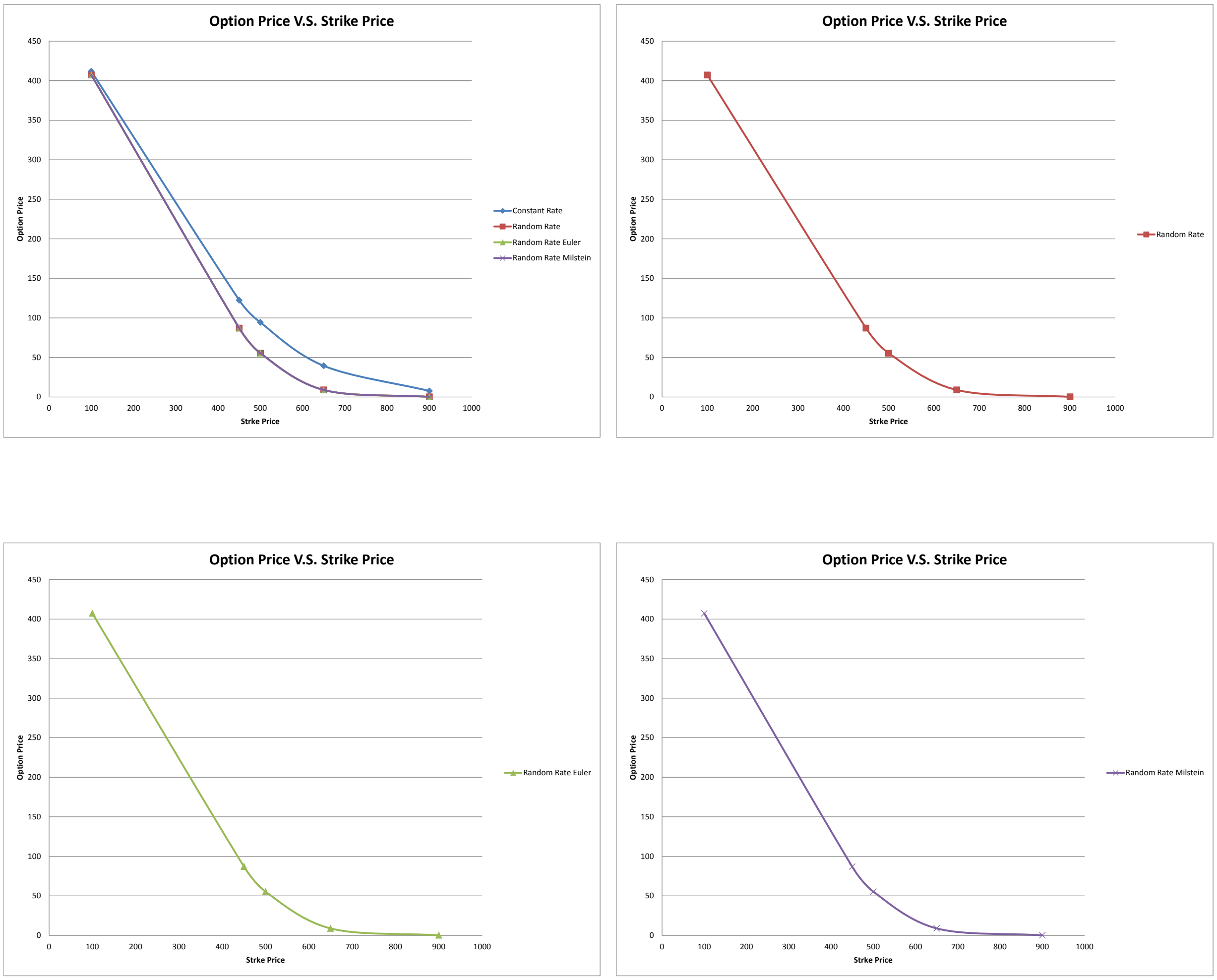

Figure 7: Option Price V.S. Strike Price, $r_0 = 0.1$

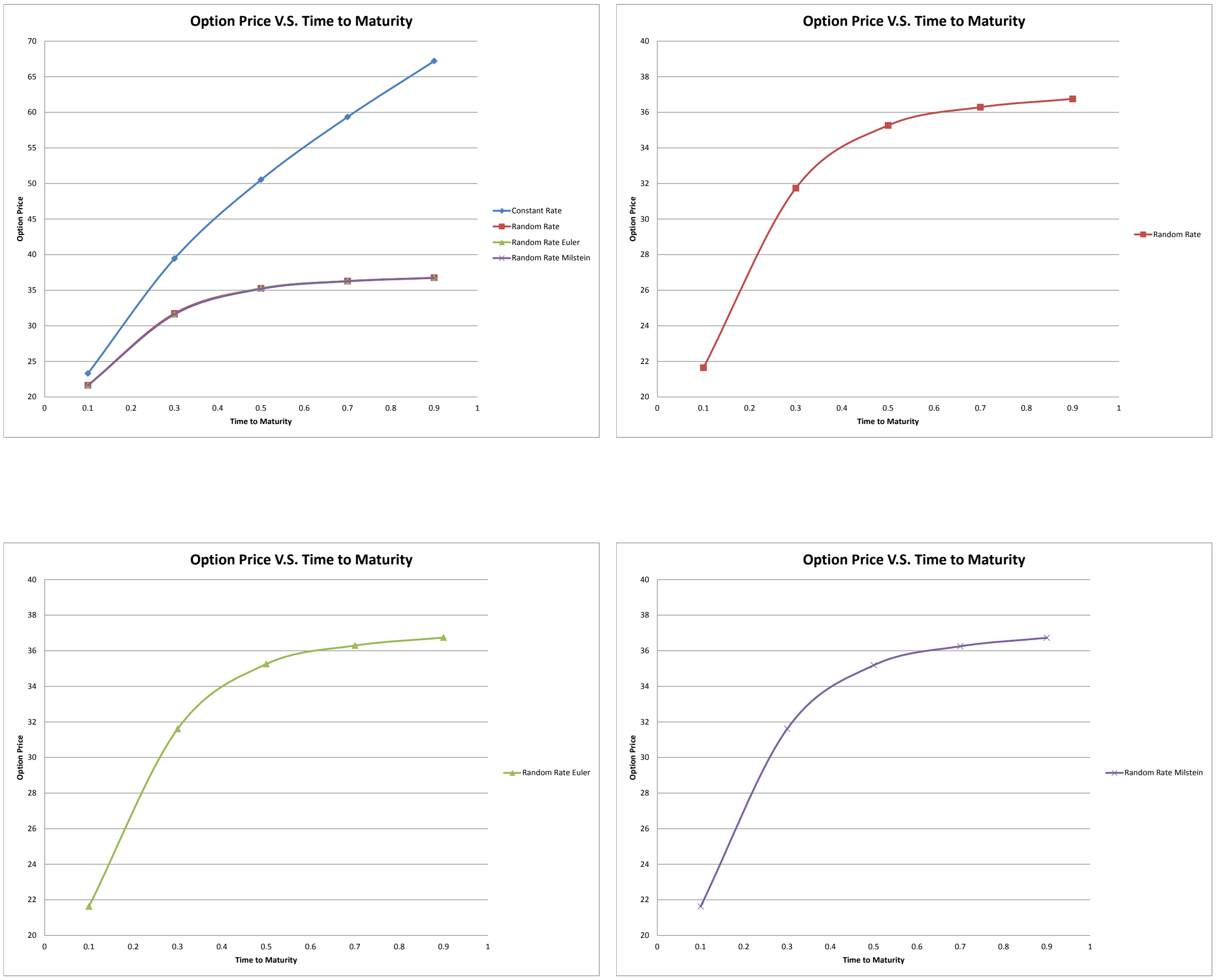

Figure 8: Option Price V.S. Time to Maturity, $r_0 = 0.0008$

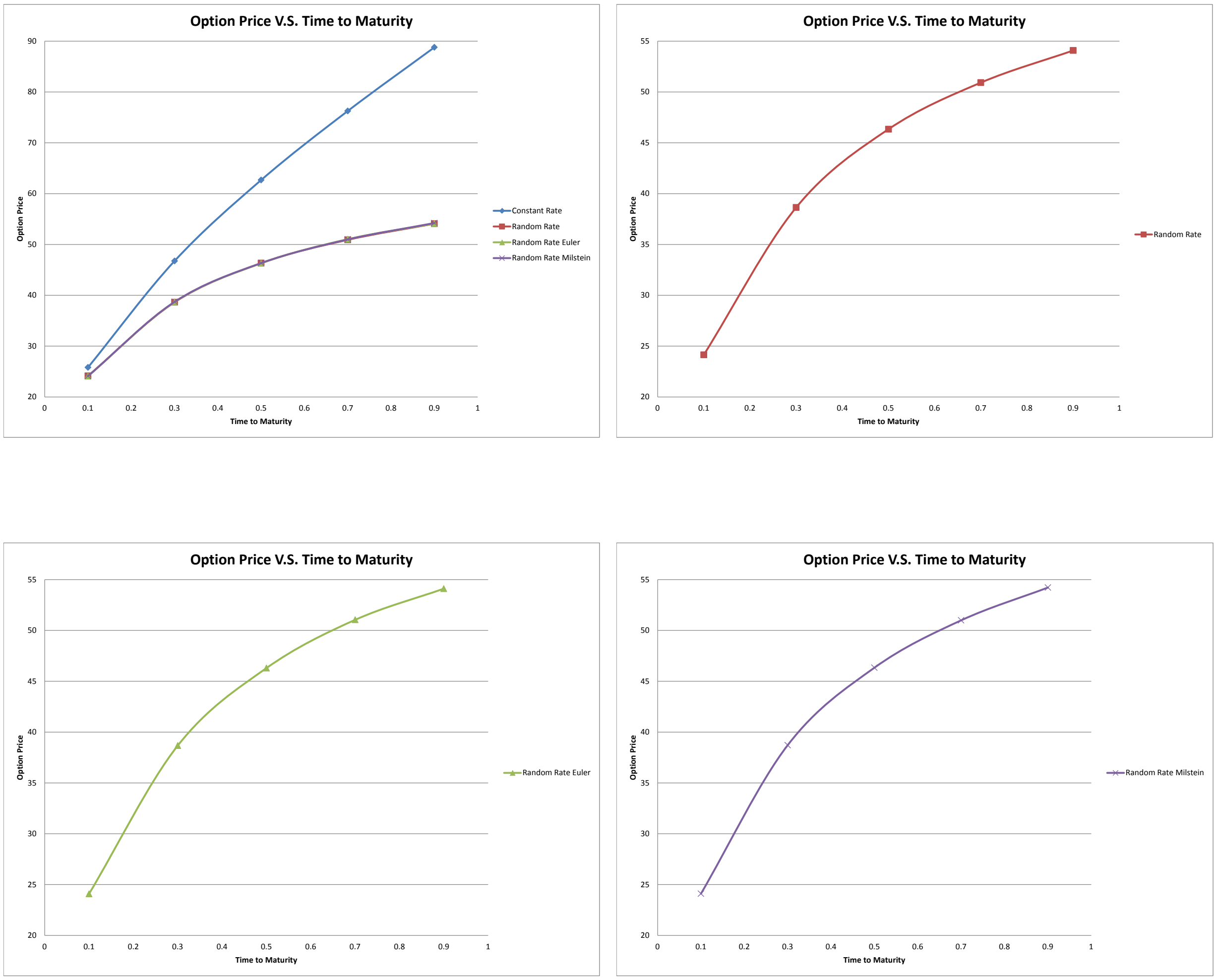

Figure 9: Option Price V.S. Time to Maturity, $r_0 = 0.1$

Bibliography

[1] D. Wilcox. Energy, futures and options: Spread options in energy markets. *New York: Goldman Saches & Co.*

[2] Geoffrey Poitra. Spread options, exchange options, and arithmetic brownian motion. *Journal of Futures Markets*, 18:487–517, 1998.

[3] W. Margrabe. The vlaue of an option to exchange one asset for another. *The journal of Finance*, 33:177–186, 1978.

[4] Deng-Feng Wang and Craig Liu. Exchange options and spread options with stochastic interest rates. 1999.

[5] Bernt Oksendal. *Stochastic Differential Equations: An Introduction with Applications.* Springer-Verlag Berlin Heidelberg New York, 6th edition, 2005.

[6] Steven E. Shreve. *Stochastic Calculus for Finance II Continuous-Time Models*. Springer, 2004.

[7] Frank J. Jones. Spreads: Tails, turtles, and all that. *The Journal of Futures Markets*, 11:565–596, 1991.

[8] Frank J. Jones. Copper-bottom pricing. *RISK*, May, 1995.

[9] New York mercantile Exchange. *Crack Spread Handbook*.

[10] Aram Sogomonian Shejie Deng, Balke Johnson. Spark spread options and the valuation of electricity generation assets. *Proceeding HICSS '99 Proceedings of the Thirty-Second Annual Hawaii International Conference on System Sciences*, 3, 1999.

[11] A. Mbanefo. Co-movement term structure and the valuation of energy spread options. pages 89–111.

[12] Mark Garman. Spread the load. *RISK*, 5, 1992.

[13] Oldrich Vasicek. An equilibrium characterisation of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.

[14] Achim Klenke. *Probability Theory A Comprehensive Course*. Springer, 2000.

[15] Jan Vecer. *Stocashtic Finance A Numeraire Aproach*. CRC Press, 2011.

[16] N.H.Bingham and Rudiger Kiesel. *Risk-Neutral Valuation Pricing and Hedging of Financial Derivatives*. Springer, 2004.

[17] J.M.Harrison and D. Kreps. Martingales and arbitrage in multi-period security markets. *Journal of Economic Theory*, 20:381–408, 1979.

[18] J.M.Harrison and S.Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11:215–260, 1981.

[19] Rick Durrett. *Probability: Theory and Examples*. Thomson, 2005.

[20] Lars Peter Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50:1029–1054, 1982.

[21] James Douglas Hamilton. *Time Series Analysis*. Princeton University Press, 1994.

[22] Laszlo Matyas. *Generalized Methods of Moments Estimation*. Princeton University Press, 1999.

[23] Francis A. Longstaff Anthony B. Sanders K.C. Chan, G. Andrew Karolyi. An empirical comparison of alternative models of the short-term interst rate. *The Journal of Finance*, 47:1209–1227, 1992.

[24] Eckhard Platen Peter E. Kloeden. *Numerical Solution of Stochastic Differential Equations*. Springer, 1999.

[25] Paul Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer, 2000.