NUMERICAL ALGORITHMS FOR STOCK OPTION VALUATION

by

Scott B. Nickleach

B.S. Mathematics, Slippery Rock University, 2002

M.A. Statistics, University of Pittsburgh, 2004

Submitted to the Graduate Faculty of

Faculty of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2008

UNIVERSITY OF PITTSBURGH

FACULTY OF ARTS AND SCIENCES

This dissertation was presented

by

Scott B. Nickleach

It was defended on

May 7th, 2008

and approved by

Leon Gleser

Leonce Bargeron

Satish Iyengar

Thomas Savits

Dissertation Director: Leon Gleser

NUMERICAL ALGORITHMS FOR STOCK OPTION VALUATION

Scott B. Nickleach, PhD

University of Pittsburgh, 2008

Since the formulation by Black, Scholes, and Merton in 1973 of the first rational option pricing formula which depended only on observable values, the volume of options traded daily on the Chicago Board of Exchange has grown rapidly. In fact, over the past three decades, options have undergone a transformation from specialized and obscure securities to ubiquitous components of the portfolios of not only large fund managers, but of ordinary individual investors. Essential ingredients of any successful modern investment strategy include the ability to generate income streams and reduce risk, as well as some level of speculation, all of which can be accomplished by effective use of options.

Naturally practitioners require an accurate method of pricing options. Furthermore, because today's market conditions evolve very rapidly, they also need to be able to obtain the price estimates quickly. This dissertation is devoted primarily to improving the efficiency of popular valuation procedures for stock options. In particular, we develop a method of simulating values of European stock options under the Heston stochastic volatility model in a fraction of the time required by the existing method. We also develop an efficient method of simulating the values of American stock option values under the same dynamic in conjunction with the Least-Squares Monte Carlo (LSM) algorithm. We attempt to improve the efficiency of the LSM algorithm by utilizing quasi-Monte Carlo techniques and spline methodology. We also consider optimal investor behavior and consider the notion of option trading as opposed to the much more commonly studied valuation problems.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

Dissertation prefaces often consist of deep self reflections. Conveyance of inner thoughts and feelings experienced throughout the years regarding the insurmountable nevertheless conquered obstacle, the next and even more gruesome obstacle, *ad infinitum*. For the reader's sake, those jottings remain confined within a thick notebook that my Mom keeps in a trunk in her attic.

I would like, however, to acknowledge a group of individuals, the contributions from whom have been invaluable in many ways. I thank Morley Davidson for initially helping to convince me that I wanted to pursue a Ph.D. I thank William Lindgren for helping to transform that want into the belief that I could actually obtain a Ph.D. This included him going well out of his way to help me eradicate what used to be an overwhelming fear of public speaking... and now I set off to pursue a career steeped in just that - what irony! And thanks to Robert Buck for helping to turn that belief into opportunity by pointing me in the direction of statistics, and also in the direction of Pitt.

In addition to the entire faculty, staff, and graduate students of the Statistics Department at Pitt, I wish to offer a special thanks to my committee members Leonce Bargeron, Satish Iyengar, and Tom Savits. Thanks to Allan Sampson for allowing me to harass him so much over the years (and for not harassing back too much), and also for offering guidance and support both in and out of the classroom. And of course Leon Gleser, whose contributions, influence, and support have had an immeasurable effect on my development over the past six years. Dr. Gleser's "open door policy" and the freedom he gave me to pursue my own research interests have been two especially appreciated aspects of our relationship.

Finally, I would like to thank my Mom, Pam, Dana, and Myron for their support, encouragement, and tolerance, both over the years and in the years to come.

1.0 BACKGROUND AND MOTIVATION

In the first two sections of this chapter, we provide the finance and applied probability backgrounds inherent to the various option pricing techniques, optimal early exercise strategies, and dynamic trading strategies examined throughout the dissertation. In the final section, we discuss motivating issues and outline the remainder of the dissertation.

1.1 FINANCE BACKGROUND

We begin by presenting a brief overview of the financial terminology used throughout the dissertation. Note that we also provide a glossary of finance terms in Appendix A. A **derivative security** or **contingent claim** is a financial asset whose value is derived in part from the value and characteristics of some other underlying asset(s). Common types of derivative securities are options, forwards, futures, and interest rate swaps. We focus on options throughout the dissertation, although much of our methodology readily applies in other settings.

A **stock option** is a derivative which gives its holder the right, but not the obligation, to engage in some future transaction involving the underlying stock. A **call (put) option** gives its holder the option to purchase (sell) a share of the underlying stock for a pre-specified price, known as the **strike price**. Option contracts also have an **expiration date**. The holder of a **European** option has the right to **exercise** the option only on the option's expiration date. The holder of an **American** option has the right to exercise the option on or at any time before the expiration date. Of course, rather than exercising an option, the holder may instead choose to simply sell the option at any time before its expiration date.

The value of an option contract depends on several parameters including the value of shares of the underlying stock, the strike price of the option, and the amount of time remaining until expiration. An option's value also depends on characteristics of the underlying **asset price dynamic**, including the volatility and the risk-free interest rate. A rational pricing theory dictates that the value of a call (put) option increases as the value of the underlying asset increases (decreases). The value of a call (put) option decreases as the strike price increases (decreases). For non-dividend paying stocks, the value of all American options and European calls increase as the amount of time until expiration increases. The value of a European put as a function of time remaining until expiration increases from zero to a time corresponding to a maximum value and then decreases beyond this time. The value of any option increases as the volatility of the underlying asset increases. The value of a call (put) option increases (decreases) as the risk-free interest rate increases. We relate changes in the values of options to changes in these underlying variables more formally in later sections.

An investor is **long** a position if he owns the asset making up the position. An investor is **short** a position if he has sold the asset making up the position, without having owned it in the first place. In this case, the investor borrows the asset and sells it with the agreement of returning the asset to the original holder sometime in the future. A position consisting of long and short positions in call (put) options on the same underlying asset, in which the options have the same expiration date but different strike prices is called a **vertical call (put) spread**. A position consisting of long and short positions in call (put) options on the same underlying asset, in which the options have different expiration dates but the same strike price is called a **horizontal call (put) spread**. A position consisting of long and short positions in call (put) options on the same underlying asset, in which the options have different expiration dates and different strike prices is called a **diagonal call (put) spread**.

Arbitrage occurs when an investor is able to earn certain profits without exposure to risk. **Hedging** is a means of reducing the exposure to risk associated with holding one asset by holding other assets whose returns are usually correlated with the first. A unit investment today which earns **continuously compounded interest at rate** r will be worth e^{rt} t years from now. In a landmark contribution to the field of option pricing, Black and Scholes [5]

were able to develop a closed form solution for the price of European options under certain conditions. Their approach relies on a no arbitrage paradigm, which stems from observing that the return on an option can be perfectly replicated by continuously rebalancing a hedged portfolio consisting of positions in shares of the underlying stock and a risk free asset which earns continuously compounded interest at a rate, r, i.e. government bonds.

An American option is always worth at least as much as the corresponding European option since the American option holder could always choose to act in the same manner as the rational holder of the European option by simply holding the option and exercising whenever "in-the-money" at expiry. In a major extension of Black and Scholes, Merton [30] shows that if the underlying stock pays no dividends, it is never optimal to exercise an American call option early, and hence that the value of the American call is the same as the value of the European call yielded by the Black-Scholes formula. With this exception, closed form solutions for values of American style options are rarely available. Black and Scholes valuation formulas obtain under fairly strict assumptions. Relaxing these assumptions generally renders closed form solutions unavailable, so that one must implement numerical techniques to find approximate values. In the following chapter, we more closely examine these contributions of Black, Scholes, and Merton, as well as some of the numerical valuation procedures that have since been developed.

1.2 APPLIED PROBABILITY BACKGROUND

Applied probability lies at the heart of much of the option pricing literature. In this section we provide a brief probability and stochastic calculus background, around which many of the ideas in later sections are centered.

For an arbitrary nonempty space Ω, a class $\mathcal{F}$ of subsets of Ω is a field if it contains Ω itself, and is closed under the formation of complements and finite unions. A class $\mathcal{F}$ of subsets of Ω is a σ-field if it is a field and is also closed under the formation of countable unions. The generated field, $\sigma(\mathcal{F})$, has the properties:

- $\mathcal{F} \subset \sigma(\mathcal{F})$

- $\sigma(\mathcal{F})$ is a σ-field
- if $\mathcal{F} \subset \mathcal{G}$ and $\mathcal{G}$ is a σ-field, then $\sigma(\mathcal{F}) \subset \mathcal{G}$

Let $\mathcal{I}$ be the class of subintervals of $\Omega = \mathbb{R}$. Then the elements in $\sigma(\mathcal{I}) = \mathcal{R}$ are the linear **Borel sets** of $\mathbb{R}$.

A set function μ on a field $\mathcal{F}$ is a **measure** if it satisfies the conditions:

- $\mu(A) \in [0, \infty]$ for $A \in \mathcal{F}$
- $\mu(\emptyset) = 0$
- if $A_1, A_2, \ldots$ is a disjoint sequence of sets in $\mathcal{F}$ and if $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$, then

$$\mu(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mu(A_k).$$

A set function P on a field $\mathcal{F}$ is a **probability measure** if it satisfies the conditions:

- $0 \leq P(A) \leq 1$ for $A \in \mathcal{F}$
- $P(\emptyset) = 0, P(\Omega) = 1$
- if $A_1, A_2, \ldots$ is a disjoint sequence of sets in $\mathcal{F}$ and if $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$, then

$$P(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k).$$

If $\mathcal{F}$ is a σ-field in Ω, $(\Omega, \mathcal{F})$ is a **measurable space**. If $\mathcal{F}$ is a σ-field in Ω and μ is a measure on $\mathcal{F}$, $(\Omega, \mathcal{F}, \mu)$ is called a **measure space**. If $\mathcal{F}$ is a σ-field in Ω and P is a probability measure on $\mathcal{F}$, $(\Omega, \mathcal{F}, P)$ is called a **probability space**. A probability space is **complete** if $A \subset B$, $B \in \mathcal{F}$, and $P(B) = 0$ together imply that $A \in \mathcal{F}$. For two measurable spaces $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$, consider the mapping $T : \Omega \to \Omega'$ and define $T^{-1}A' = [\omega \in \Omega : T\omega \in A']$ for $A' \in \Omega'$. T is **measurable** if $T^{-1}A' \in \mathcal{F}$ for each $A' \in \mathcal{F}'$. A real function f on Ω is measurable (or measurable $\mathcal{F}$) if $f^{-1}A = [\omega : f(\omega) \in A] \in \mathcal{F}$ for every $A \in \mathcal{R}$.

For a nonnegative real measurable function f and for each finite decomposition $\{A_i\}$ of Ω into $\mathcal{F}$-sets, consider the sums

$$\sum_i [\inf_{\omega \in A_i} f(\omega)] \mu(A_i).$$

The integral of f with respect to μ is defined as

$$\int f d\mu = \sup \sum_i [\inf_{\omega \in A_i} f(\omega)] \mu(A_i),$$

where the supremum extends over all finite decompositions $\{A_i\}$ of Ω into $\mathcal{F}$-sets. A general measurable function f is **integrable** if $\int f^+ d\mu$ and $\int f^- d\mu$ are both finite.

A **random variable** is a real measurable function X on a probability space $(\Omega, \mathcal{F}, P)$. The **distribution** of the random variable is the probability measure μ defined by

$$\mu(A) = P[X \in A], \;\; A \in \mathcal{R}.$$

The **characteristic function** of a random variable X with distribution μ is

$$\eta(t) = E[e^{itX}] = \int_{-\infty}^{\infty} e^{itx} \mu(dx).$$

Because different measures cannot share the same characteristic function, it is possible to recover a measure μ by *inverting* the corresponding characteristic function. In particular,

$$\mu(a, b] = \lim_{T \to \infty} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \eta(t) dt.$$

Let X be an integrable random variable on $(\Omega, \mathcal{F}, P)$, and let $\mathcal{G}$ be a σ-field contained in $\mathcal{F}$. There exists a random variable $E(X||\mathcal{G})$, having the properties

- $E(X|\mathcal{G})$ is measurable $\mathcal{G}$ and integrable
- $E(X|\mathcal{G})$ satisfies

$$\int_G E(X|\mathcal{G}) dP = \int_G X dP, \;\; G \in \mathcal{G}.$$

$E(X|\mathcal{G})$ is called the **conditional expectation** of X given $\mathcal{G}$. In general there will be many such random variables, any one of which is called a **version** of the conditional expectation; any two versions are equal on a set of probability 1.

Let T represent a collection of time points. A **stochastic process** $(X(t), t \in T) = (X(t,\omega), t \in T, \omega \in \Omega)$ is a collection of random variables defined on Ω. For our purposes, T will either be an interval $[a, b]$ or a sequence of discrete time points. Assume that $(\mathcal{F}_t, t \geq 0)$ is a collection of σ-fields all in the same space Ω, and that each $\mathcal{F}_t$ is a subset of a larger σ-field $\mathcal{F}$, also in Ω. The collection $(\mathcal{F}_t, t \geq 0)$ is called a **filtration** if $\mathcal{F}_s \subset \mathcal{F}_t$ for all $0 \leq s \leq t$. The stochastic process $(X(t), t \in T)$ is **adapted to the filtration** $(\mathcal{F}_t, t \geq 0)$ if $\sigma(X(t)) \subset \mathcal{F}_t$ for all $t \geq 0$.

A stochastic process $(X(t), t \geq 0)$ is a **Markov process in continuous time** if for all $k \geq 1, 0 \leq t_1 \leq \cdots \leq t_k \leq u$ and $H \in \mathcal{R}$, $P[X(u) \in H|X(t_1), \ldots, X(t_k)] = P[X(u) \in H|X(t_k)]$. A stochastic process $(X(n), n = 0, 1, \ldots)$ is a **Markov process in discrete time** if $P[X(n+1) \in H|X(0), \ldots, X(k)] = P[X(n+1) \in H|X(k)]$.

A stochastic process $(X(t), t \geq 0)$ is a **continuous-time martingale with respect to the filtration** $(\mathcal{F}_t, t \geq 0)$ if the following hold:

- $E|X(t)| < \infty$ for all $t \geq 0$
- $X(t)$ is measurable $\mathcal{F}_t$
- $E(X(t)|\mathcal{F}_s) = X(s) \ \forall \ 0 \leq s \leq t$

A stochastic process $(X(n), n = 0, 1, \ldots)$ is a **discrete-time martingale with respect to the filtration** $(\mathcal{F}_n, n = 0, 1, \ldots)$ if the following hold:

- $E|X(n)| < \infty$ for all $n = 0, 1, \ldots$
- $X(n)$ is measurable $\mathcal{F}_n$
- $E(X(n+1)|\mathcal{F}_n) = X(n) \ \forall \ n = 0, 1, \ldots$

If we replace the final conditions in the two previous definitions by $E(X(t)|\mathcal{F}_s) \geq X(s)$ and $E(X(n+1)|\mathcal{F}_n) \geq X(n)$, then $(X(t), t \geq 0)$ is a **continuous-time submartingale** and $(X(n), n = 0, 1, \ldots)$ is a **discrete-time submartingale**. If we instead replace the conditions with $E(X(t)|\mathcal{F}_s) \leq X(s)$ and $E(X(n+1)|\mathcal{F}_n) \leq X(n)$, then $(X(t), t \geq 0)$ is a **continuous-time supermartingale** and $(X(n), n = 0, 1, \ldots)$ is a **discrete-time supermartingale**.

Let $\Pi = \{t_0, t_1, \ldots, t_n\}$ be a finite partition of $[0, T]$, i.e. $0 = t_0 \leq t_1 \leq \cdots \leq t_n = T$. Define the maximum step size of the partition as $||\Pi|| = \max_{j=0,\ldots,n-1}(t_{j+1} - t_j)$. The **quadratic variation** up to time t of a sufficiently regular function $f(t)$ defined on $[0, T]$ is

$$\lim_{||\Pi||\to 0} \sum_{j=0}^{n-1} [f(t_{j+1}) - f(t_j)]^2,$$

where $0 = t_0 < t_1 < \cdots < t_n = T$.

A stochastic process $(W(t), t \geq 0)$ is called **Brownian motion** or a **Wiener process** if the following hold:

- $W(0) = 0$
- $W(t) - W(s)$ follows a Normal distribution with mean 0 and variance $\sigma^2(t - s)$ for $0 \leq s \leq t$.
- $W(t_2) - W(t_1), \ldots, W(t_n) - W(t_{n-1})$ are independent for $t_1 \leq \ldots \leq t_n$.

The process $Y(t) = \mu t + \sigma W(t)$ is called **Brownian motion with drift** or **arithmetic Brownian motion**. The process $S(t) = e^{\mu t + \sigma W(t)}$, used to model many financial assets, is called **geometric Brownian motion**.

Again let $\{t_0, t_1, \ldots, t_n\}$ be a partition of $[0, T]$, with $0 = t_0 \leq t_1 \leq \cdots \leq t_n = T$. Define a process $\Delta_n(t)$ to be a **simple process** if it is constant on each subinterval $[t_j, t_{j+1})$. Then the quantity

$$I(T) = \sum_{j=0}^{n-1} \Delta_n(t_j)[W(t_{j+1}) - W(t_j)]$$

is the **Ito integral** $\int_0^T \Delta_n(t) dW(t)$ of $\Delta_n(t)$ over the interval $(0, T)$.

To define the Ito integral $\int_0^T \Delta(t) dW(t)$ for the general integrand $\Delta(t)$, we begin by assuming that $(\Delta(t), t \geq 0)$, is adapted to the Brownian filtration $\mathcal{F}_t$, and also that

$$E \int_0^T \Delta^2(t) dt < \infty.$$

It is possible to choose a sequence $\Delta_n(t)$ of simple processes such that as $n \to \infty$ the processes converge to the continuously varying $\Delta(t)$ in the sense that

$$\lim_{n\to\infty} E\int_0^T |\Delta_n(t) - \Delta(t)|^2 dt = 0.$$

The Ito integral for $\Delta(t)$ over the interval $(0, T)$ is defined as

$$\int_0^T \Delta(t)dW(t) = \lim_{n\to\infty}\int_0^T \Delta_n(t)dW(t).$$

Let $(W(t), t \geq 0)$ be a Brownian motion and let $(\mathcal{F}_t, t \geq 0)$ be an associated filtration. Let $(\Delta(t), t \geq 0)$ and $(\Theta(t), t \geq 0)$ be adapted processes. A process $(X(t), t \geq 0)$ that can be represented as

$$X(t) = X(0) + \int_0^t \Delta(u)dW(u) + \int_0^t \Theta(u)du$$

is called an **Ito process**. In terms of differentials, an Ito process can be written as

$$dX(t) = \Delta(t)dW(t) + \Theta(t)dt. \tag{1.1}$$

Let $(X(t), t \geq 0)$, be an Ito process and let $f(t, x)$ have continuous second partial derivatives. **Ito's Lemma** states that for $T \geq 0$,

$$\begin{aligned} f(X(T), T) - f(X(0), 0) = &\int_0^T f_t(X(t), t)dt + \int_0^T f_x(X(t), t)\Delta(t)dW(t) \\ &+ \int_0^T f_x(X(t), t)\Theta(t)dt + \frac{1}{2}\int_0^T f_{xx}(X(t), t)\Delta^2(t)dt, \end{aligned}$$

where the subscripts indicate the direction(s) of partial derivatives of f. In terms of differentials, Ito's Lemma can be stated as

$$df(X(t), t) = f_t(X(t), t)dt + f_x(X(t), t)dX(t) + \frac{1}{2}f_{xx}(X(t), t)dX(t)dX(t). \tag{1.2}$$

The quadratic variation of Brownian motion over the interval $(0, t)$ is t. That is to say that Brownian motion accumulates quadratic variation at a rate one per unit time. The

quadratic variation of arithmetic Brownian motion over the interval $(0,t)$ is $\sigma^2 t$. More generally, the quadratic variation of an Ito process over $(0,t)$ is $\int_0^t \Delta^2(u)du$.

A **stochastic differential equation** is an equation of the form

$$dX(u) = \beta(X(u),u)du + \gamma(X(u),u)dW(u).$$

The **Feynman-Kac Theorem** relates stochastic differential equations with partial differential equations. Let h be a Borel-measurable function and r be constant, and assume that $E[h(X(T))|X(t) = x(t)] < \infty$ for all t. Fix $T > 0$ and define the function

$$f(x,t) = e^{-r(T-t)}E[h(X(T))|\mathcal{F}_t] \text{ for } t \in [0,T].$$

Then $f(x,t)$ satisfies the partial differential equation

$$f_t(x,t) + \beta(x,t)f_x(x,t) + \frac{1}{2}\gamma^2(x,t)f_{xx}(x,t) = rf(x,t), \tag{1.3}$$

subject to the boundary condition $f(x,T) = h(x)$ for all x.

For a measurable space $(\Omega, \mathcal{F})$, consider two probability measures P and $\tilde{P}$ on Ω. We say that P is **absolutely continuous with respect to** $\tilde{P}$, written $P << \tilde{P}$, if $\forall A \in \mathcal{F}$, $\tilde{P}(A) = 0$ implies that $P(A) = 0$. We say that P and $\tilde{P}$ are **equivalent** probability measures if $P << \tilde{P}$ and $\tilde{P} << P$. The **Radon-Nikodym Theorem** says that for two probability measures P and $\tilde{P}$, $\tilde{P} << P$ holds if and only if there exists an almost surely positive random variable Z such that $E(Z) = 1$ and

$$\tilde{P}(A) = \int_A Z(\omega)dP(\omega) \quad \forall A \in \mathcal{F}. \tag{1.4}$$

If this relation holds with Z replaced by another random variable Z', then $P[Z \neq Z'] = 0$. The random variable Z is called a version of the **Radon-Nikodym derivative of** $\tilde{P}$ **with respect to** P.

Let $W(t)$, $0 \leq t \leq T$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ and let $(\mathcal{F}_t, t \geq 0)$ be an associated filtration. Let $\Theta(t), 0 \leq t \leq T$, be an adapted process. Define

$$Z(t) = \exp\{-\int_0^t \Theta(u)dW(u) - \frac{1}{2}\int_0^t \Theta^2(u)du\},$$

$$\tilde{W}(t) = W(t) + \int_0^t \Theta(u)du.$$

Set $Z = Z(T)$. **Girsanov's Theorem** states that $E(Z) = 1$, and that under the probability measure $\tilde{P}$ in 1.4, the process $\tilde{W}(t)$, $0 \leq t \leq T$, is a Brownian motion. The measure $\tilde{P}$ is called an **equivalent martingale measure**. In the finance literature, $\tilde{P}$ is often referred to as a **risk neutral measure**. Under a risk neutral measure, the discounted value of an asset price is a martingale. This is to say that under the risk neutral measure, an investor who is risk neutral is indifferent between holding a risky asset and a risk free asset. A market is **arbitrage-free** if and only if there exists a risk neutral measure. A market is **complete** if and only if there exists a unique risk neutral measure.

1.3 MOTIVATION

What is an option and why should we care about them? In one of the most celebrated contributions to the field in the past half century, Robert Merton [30] begins by saying that,

> Because options are specialized and relatively unimportant financial securities, the amount of time and space devoted to the development of a pricing theory might be questioned.

Now that nearly ten million contracts trade hands daily on the Chicago Board of Exchange and one Nobel Prize later, Merton may have a different perspective on the importance of options and option valuation. In fact, over the past three decades, options have undergone a transformation from specialized and obscure securities to ubiquitous components of the portfolios of not only large fund managers, but of ordinary individual investors. Essential ingredients of any successful modern investment strategy include the ability to generate income streams and reduce risk, as well as some level of speculation, all of which can be accomplished by effective use of options. For these reasons, practitioners require an accurate method of pricing options. Furthermore, because today's market conditions evolve very rapidly, they also need to be able to obtain the price estimates in a short period of time. It is our goal to address each of these issues.

The remainder of the dissertation is organized as follows. In Chapter 2, we summarize how the field of option valuation has evolved over the past three decades, beginning with the celebrated works of Black, Scholes, and Merton. We include overviews of several popular numerical valuation techniques for American style options. In Chapter 3, we focus on Heston's stochastic volatility model [21] and the exact simulation methods of Broadie and Kaya [8] for Heston's model and several other commonly used stochastic volatility models. We present a quasi-Monte Carlo method that yields European option price estimates in far less computational time than that required by their simulation algorithm. We also develop an alternative simulation method that again achieves the goal of yielding European option price estimates in far less time than the current methods. In Chapter 4 we discuss the well known American option valuation algorithm of Longstaff and Schwartz [29], known as the Least Squares Monte Carlo algorithm. We attempt to improve the algorithm using quasi-Monte Carlo techniques and splines. We also extend the methods of the previous chapter to American style valuation. We check the sensitivity of the algorithm to the choice of basis functions used in conjunction with non-Markovian price dynamics. In Chapter 5 we present comparisons of optimal behavior under different assumed underlying models and propose a variation of the Least Squares Monte Carlo algorithm, the goal of which is to construct optimal trading strategies rather than optimal early exercise strategies.

2.0 LITERATURE OVERVIEW

In this chapter, we summarize the major developments in the field of option valuation throughout the past three decades. We begin with a somewhat detailed rendition of the landmark contributions of Black and Scholes [5], and Merton [30]. Next we shift from diffusion models to the jump-diffusion models of Merton [31] and Cox and Ross [11] which allow for discontinuous asset price paths. This is followed with the work of Harrison and Kreps [17], and Harrison and Pliska [18],[19], which establishes the link between arbitrage-free pricing and martingales. We conclude the chapter by presenting some popular numerical valuation techniques, including finite difference methods of Schwartz [37], lattice approaches of Cox, Ross, and Rubenstein [12], and Monte Carlo simulation approaches of Boyle [6]. Note that we have provided a glossary in Appendix A and have also done our best to unify 30 years of constantly evolving notation, the most common of which can be found in Appendix B.

2.1 BLACK, SCHOLES, AND MERTON

The most celebrated and well-known work in the field of option pricing belongs to Fisher Black and Myron Scholes [5], and Robert Merton [30]. For these landmark contributions during the early 1970's, Scholes and Merton were awarded the Nobel Prize for Economics in 1997.

We begin by presenting the underlying assumptions of Black and Scholes, under which their famous option valuation formula obtains:

- The short term risk free interest rate, r, is a known constant.

- Stock price follows geometric Brownian motion with known constant variance rate σ^2.
- The underlying stock pays no dividends.
- The option is European.
- The are no transaction costs associated with trading.
- One may borrow without limit at the short term risk free interest rate.
- Short selling is allowed and the seller may use all proceeds generated by the short sale.

The main idea governing the development of the option pricing formula is that the return of an option can be perfectly replicated by continuously rebalancing a portfolio consisting of shares of the underlying stock and a risk free asset, i.e. a government bond. The return on this hedged portfolio is shown to be independent of price movements of the stock, and thus depends only on time and the values of other constants assumed to be known. This deterministic return cannot be greater than the return of the initial investment in the hedged portfolio, compounded at the risk free interest rate over the life of the option. Otherwise traders could generate arbitrage profits by borrowing at the risk free rate and using those borrowed funds to establish a position in the higher yielding hedged portfolio, (which would in turn force the yield to the equilibrium risk free rate.) For similar reasons, the deterministic return cannot be less than the risk free rate and hence must equal the risk free rate.

More specifically, let $u(s,t)$ represent the value of the option as a function of the underlying stock price s and time t. Then the number of shares of stock that should be purchased to hedge against one short option over the next brief period of time is $u_s = \partial u/\partial s$.

To convey the intuition, consider an option with $u_s = .5$. Suppose that, with time held fixed, the value of the underlying share of stock increases by 1 unit. Then the value of this portfolio consisting of one short option and the prescribed .5 shares of the underlying stock will change by approximately -1(.5) + .5(1) $= -u_s + u_s = 0$. The same is true if we consider a 1 unit decrease in the value of the underlying, so that the return is independent of the direction of the underlying price movement.[1]

Equivalently, the number of options to be shorted against one long share of the underlying stock is $1/u_s$. Then the initial equity in the hedged portfolio is $s - u/u_s$, and the change in

[1]The entire hedged portfolio at any point in time consists of $u_s(s,t)$ shares of the underlying stock and $u(s,t) - u_s(s,t)s(t)$ invested in a risk-free bond.

value of equity during a short interval dt is

$$ds - du/u_s. \tag{2.1}$$

Using differential notation, our underlying dynamic can be written as the Ito process

$$ds = \mu s dt + \sigma s dW(t),$$

where μ is the expected rate of return on the stock. Then according to Ito's Lemma 1.2, we can write

$$du = u_s ds + \frac{1}{2} u_{ss} dsds + u_t dt. \tag{2.2}$$

Recall that Brownian motion accumulates quadratic variation at a rate 1 per unit time, which can be represented using differentials as $dW(t)dW(t) = dt$. Furthermore, it can be shown that $dW(t)dt = dtdW(t) = 0$ and $dtdt = 0$. This yields

$$dsds = \mu^2 s^2 dtdt + 2\mu s^2 dW(t)dt + \sigma^2 s^2 dW(t)dW(t) = \sigma^2 s^2 dt,$$

and from Equation 2.2, we obtain

$$du = u_s ds + \frac{1}{2} u_{ss} \sigma^2 s^2 dt + u_t dt. \tag{2.3}$$

Substituting equation 2.3 into equation 2.1, the change in value of equity in the hedged portfolio is

$$-(\frac{1}{2} u_{ss} \sigma^2 s^2 + u_t) dt/u_s. \tag{2.4}$$

But the return on the equity in the hedged position must be at the risk free rate, rdt. This allows us to equate 2.1 and 2.4 as

$$-(\frac{1}{2} u_{ss} \sigma^2 s^2 + u_t) dt/u_s = (ds - du/u_s) r dt. \tag{2.5}$$

Rearranging 2.5 yields the following partial differential equation which governs the value of the option[2]:

$$u_t = ru - rsu_s - \frac{1}{2}\sigma^2 s^2 u_{ss}. \tag{2.6}$$

For a call option with strike price K, on the expiration date T, the value of the option satisfies $u(s,T) = \max(s-K, 0)$. The solution to 2.6, subject to this boundary condition, is the call option valuation formula:

$$u(s,t) = s\Phi(c_1) - Ke^{r(t-T)}\Phi(c_2) \tag{2.7}$$

$$c_1 = \frac{\ln(s/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$
$$c_2 = \frac{\ln(s/K) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

where $\Phi(\cdot)$ is the cumulative standard normal distribution function. Note that μ does not appear in 2.7.

Letting $u_{put}(s,t)$ be the value of a European put, a minor modification to the boundary condition in the above derivation yields

$$u_{put}(s,t) = -s\Phi(c_1) + Ke^{r(t-T)}\Phi(-c_2). \tag{2.8}$$

Note that in the valuation formulas, the factor $(T-t)$ is always multiplied by r or σ^2. Thus an increase in time until expiration has the same effect on the value of the option as the same percentage increase in both r and σ^2. The partial derivatives of the valuation formula are often referred to as an option's "Greeks". The quantity $u_s = \partial u/\partial s$ is commonly referred to as **delta** and $u_t = \partial u/\partial t$ is referred to as **theta**. Other quantities often used in hedging under more general dynamics are an option's **rho**, **vega**, and **gamma**, which are defined as $\partial u/\partial r$, $\partial u/\partial \sigma^2$, and $\partial^2 u/\partial s^2$, respectively. Figure 1 illustrates the value of a European call option with $K = 15$, $r = .05$, and $\sigma = .25$. Note the boundary condition at $t = 0$.

[2]We use the Feynman-Kac theorem to arrive at the same partial differential equation in Section 2.2.

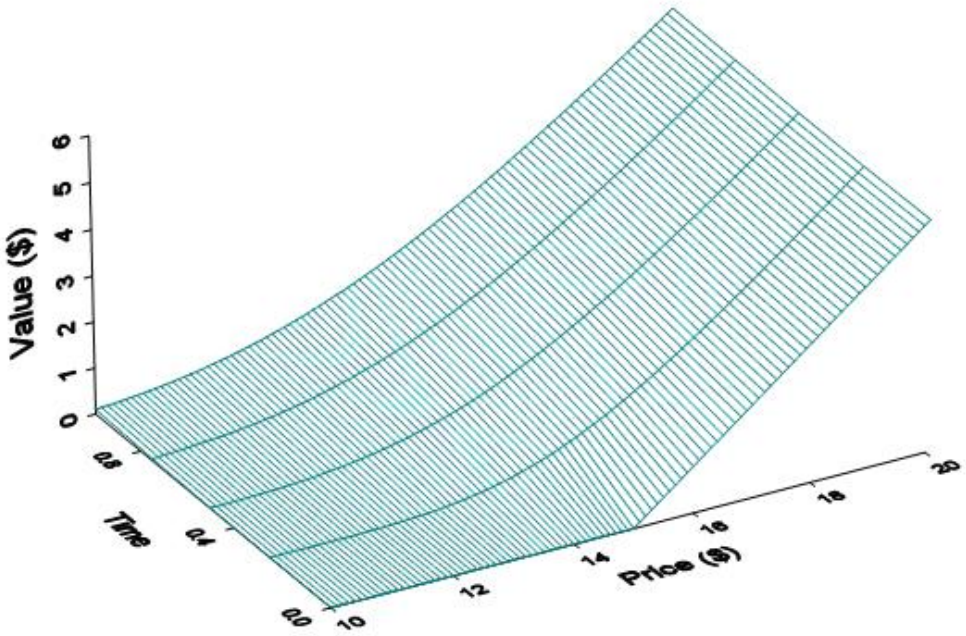

Figure 1: The value of a European call option.

Merton [30] extended the work of Black and Scholes in various directions. Under the same assumptions, he shows that the value of a call option increases continuously as any of T, r, or σ^2 increases, in each case approaching from below the stock price s. He demonstrates using a convexity argument that an option on a basket of stocks is less valuable than an individual option on each of the stocks in the basket, which should dissuade the option buyer from diversification. He also shows that for non-dividend paying stocks, one should never exercise an American call option prior to its expiration. Thus the early exercise feature of the American option is worthless and the option must have the same value as the corresponding European call option. This relationship, however, is not necessarily true for call options on dividend paying stocks or for put options. For (discrete) dividend paying stocks, he shows that an optimal early exercise strategy consists of exercising the option just before the dividend is paid, whenever the option is in-the-money and the amount of the dividend is greater than the time value component of the option's value just before the dividend is

paid. We examine further contributions of Merton in the following section.

2.2 DISCONTINUOUS ASSET PRICE PATHS

Many economists believe in exponential growth and are thus content using geometric Brownian motion as a price model. However, much empirical evidence exists indicating that stocks do not have continuous price paths. For an exposition supporting discontinuous stock price paths, see Akgiray and Booth [2]. In this section, we examine European option pricing techniques which utilize these discontinuous price dynamics.

Two critical assumptions in Black and Scholes are that assets have continuous sample paths and that trading takes place continuously. In reality, trading does not take place continuously; however, it is shown in Merton and Samuelson [32] that the Black and Scholes continuous trading solution is a valid asymptotic approximation to the discrete trading solution. Their approach fails, however, when the asset price dynamic doesn't generate continuous sample paths.

Merton [31] and Cox and Ross [11] assume that the total change in a stock price consists of two components. Regular fluctuations in price due to temporary imbalance between supply and demand, changes in the economic outlook, etc. can be modeled using a standard geometric Brownian motion. Irregular fluctuations in price due to the arrival of important information such as earnings reports, company issued guidance, etc. can be modeled using a Poisson-driven jump process. The Poisson distributed event is the arrival of an important piece of information. Arrivals are assumed to be independently and identically distributed, and the probability of an event occurring during a time interval dt is λdt. Whenever the Poisson event occurs, there is a drawing from some distribution to determine the effect, ζ, of the information on the price of the stock. The stock price returns can be written as

$$ds/s = (\alpha - \lambda k)dt + \sigma dw + dq \tag{2.9}$$

where α is the instantaneous expected return on the stock, σ^2 is the instantaneous variance of the return on the stock, conditional on no new important information arriving, dw is

standard Brownian motion, $q(t)$ is the independent Poisson process, dq is independent of dw , λ is the mean number of arrivals per unit time, and $k = E(\zeta - 1)$, where $(\zeta - 1)$ is the random percentage change if the Poisson event occurs. Under this dynamic, $(\alpha - \lambda k)dt$ models the deterministic part of the return, σdw models the regular price fluctuations, and dq models the irregular price fluctuations. Note that the special case $\lambda = 0$ corresponds to the Black-Scholes setting. If α, λ, k, and σ are constants, then

$$S(t) = s(0) \exp[(\alpha - \frac{\sigma^2}{2} - \lambda k) + \sigma Z(t)]\zeta(N).$$

Here

$$\zeta(N) = \begin{cases} 1, & \text{if } N = 0 \\ \prod_{j=1}^{N} \zeta_j, & \text{if } N \geq 1. \end{cases}$$

where the ζ_j are independent and identically distributed, $Z(t) \sim N(0, t)$, and $N \sim \text{Poisson}(\lambda t)$.

Now suppose that an option price u, can be written as a function of s and t, and that $u = u(s, t)$ is twice continuously differentiable. Then

$$du/u = (\alpha_u - \lambda k_u)dt + \sigma_u dw + dq_u \tag{2.10}$$

where α_u is the instantaneous expected return on the option, σ_u^2 is the instantaneous variance of the return on the option, conditional on no Poisson event occurring, $q_u(t)$ is an independent Poisson process with parameter λ, and $k_u = E(\zeta_u - 1)$. Using Ito's lemma, we also have the relationships:

$$\alpha_u = [\frac{1}{2}\sigma^2 s^2 u_{ss}(s, t) + (\alpha - \lambda k)s u_s(s, t) + u_t + \lambda E[u(s\zeta, t) - u(s, t)]]/u(s, t) \tag{2.11}$$

$$\sigma_u = u_s(s, t)\sigma s/u(s, t).$$

Now consider a portfolio consisting of a stock, an option on the stock, and a risk-free asset, in proportions p_1, p_2, and p_3, where $\sum p_j = 1$. If π denotes the value of the portfolio, the return on the portfolio can be written as

$$d\pi/\pi = (\alpha_\pi - \lambda k_\pi)dt + \sigma_\pi dw + dq_\pi \tag{2.12}$$

where terms with a π subscript have meanings analogous to those above with a u subscript. Then from 2.9 and 2.10, we have

$$\alpha_\pi = p_1(\alpha - r) + p_2(\alpha_u - r) + r \tag{2.13}$$

$$\sigma_\pi = p_1\sigma + p_2\sigma_u$$

$$\zeta_\pi - 1 = p_1(\zeta - 1) + p_2[u(s\zeta, t) - u(s, t)]/u(s, t).$$

In the Black-Scholes setting where $\lambda = 0$, we could eliminate the risk of this portfolio return by choosing $p_1 = p_1^*$ and $p_2 = p_2^*$, so that $p_1^*\sigma + p_2^*\sigma_u = 0$. Using these weights assures that the return on the portfolio will equal r if we are to avoid arbitrage. However in the presence of jumps, using the weights p_1^* and p_2^* will not eliminate risk. Furthermore, there are no weights p_1 and p_2 which will eliminate the jump risk. Thus, an alternative approach to deriving an option pricing formula is required.

Consider the return characteristics of a portfolio when using the Black-Scholes hedge proportions, p_1^* and p_2^*:

$$d\pi^*/\pi^* = (\alpha_\pi^* - \lambda k_\pi^*)dt + \sigma_\pi dw + dq_\pi^* \tag{2.14}$$

This return is a pure jump process since we have hedged out the continuous movements in the stock and option price. Equation 2.14 indicates that the mean return on the portfolio will usually be $(\alpha_\pi^* - \lambda k_\pi^*)$, but that on average, once every $1/\lambda$ units of time, the portfolio value will jump.

If the cause of the jumps is firm-specific information, then the jump component of the stock's return should be uncorrelated with the rest of the market. Under the p^* hedge

portfolio, the only uncertainty in the return is the jump component, and hence the correlation of returns of this portfolio with those of the market under the Capital Asset Pricing model (see, for example, Sharpe [38] and Lintner [28]) is zero, and the expected return must equal the riskless rate, r. Hence $\alpha_\pi^* = r$, and 2.13 implies $p_1^*(\alpha - r) + p_2^*(\alpha_u - r) = 0$, or

$$(\alpha - r)\sigma = (\alpha_u - r)/\sigma_u \tag{2.15}$$

Now 2.11 and 2.14 imply

$$\frac{1}{2}\sigma^2 s^2 u_{ss}(s,t) + (r - \lambda k)s u_s(s,t) + u_t(0,t) + \lambda E[u(s\zeta, t) - u(s,t)] = 0 \tag{2.16}$$

subject to the boundary conditions $u(0,t) = 0$ and $u(s,0) = \max(0, s - K)$. Although a closed-form solution to 2.16 cannot be obtained without further specification of the distribution of ζ, a form which lends to relatively simple computational approximations is available.

Recall the Black-Scholes call option pricing formula 2.7. Define the random variable $\tilde{\zeta}_N$ to have the same distribution as the product of N i.i.d. random variables, each of which is distributed identically to the random variable ζ from 2.9. Let $\tilde{\zeta}_0 = 1$, and define E_N as the expectation operator under the distribution of $\tilde{\zeta}_N$. Then the solution to 2.16 can be written as

$$u(s,t) = \sum_{N=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^N}{N} [E_N[u(S(T)\tilde{\zeta}_N e^{-\lambda k t}, T)]] \tag{2.17}$$

As well as considering discontinuous asset price paths similar to those presented above, Cox and Ross [11] go on to specify an entirely general option pricing formula, based on the idea of risk neutral investor preferences. In an investment environment where risk neutrality prevails, the expected return on the stock and the option must both equal the risk free rate, r. Then

$$E[\frac{S(T)}{s(t)} | s(t)] = e^{r(T-t)}. \tag{2.18}$$

Similarly, for a general European option with boundary value $u(s,T) = h(s)$,

$$E[\frac{u(S(T),T)}{u(s(t),t)}|s(t)] = \frac{1}{u(s(t),t)}E[h(s(T))|s(t)] = e^{r(T-t)} \tag{2.19}$$

or

$$u(s,t) = e^{-r(T-t)}E[h(S(T))|s(t)] = e^{-r(T-t)}\int h(S(T))dF(s(T)|s(t)), \tag{2.20}$$

where F is the terminal distribution of the stock price, given its current value. This formula states that the correct value of the option in a risk neutral world is the discounted expected value of the payoff function under a particular probability measure.

As an application of the general theory, consider the setting of Section 2.1. Girsanov's theorem implies the risk neutral dynamic can be written as

$$ds = rsdt + \sigma sdW(t). \tag{2.21}$$

Taking $\beta = rs$, $\gamma = \sigma s$, and $h(s) = u(s,T) = \max(s-K,0)$, the Feynman-Kac theorem 1.3 asserts that $u(s,t) = e^{-r(T-t)}E[h(S(T))|s(t)]$ satisfies

$$u_t + rsu_s + \frac{1}{2}\sigma^2 s^2 u_{ss} = ru$$

subject to the boundary condition h, which is the same as 2.6.

2.3 RELATIONSHIP BETWEEN RISK NEUTRALITY AND MARTINGALES

The section focuses on the connections between asset pricing and probability theory established by Harrison and Kreps in [17], and Harrison and Pliska in [18],[19], which provide a formal basis for Equation 2.20. We shall also convey the intuition behind the discrete time versions of the First and Second Fundamental Theorems of Asset Pricing.

Specify and fix a probability space $(\Omega, \mathcal{F}, P)$. Also specify a time horizon T, the date before which all economic activity under consideration transpires, and a filtration $\mathbb{F} =$

$\{\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_T\}$, where $\mathcal{F}_0 \subseteq \ldots \subseteq \mathcal{F}_T$. Here, Ω is the set of all possible states of the economic environment, and $\mathcal{F}_t$ is the σ-field of distinguishable events at time t.

Consider an $m+1$ dimensional stochastic process $\hat{S} = (\hat{S}(t), t = 0, 1, \ldots, T)$ with component processes $S^{(0)}, S^{(1)}, \ldots$, and $S^{(m)}$. Each component process is adapted to $\mathbb{F}$. Interpret $S^{(j)}(t)$ as the price at time t of asset j, and think of $S^{(0)}(t)$ as the price process of a risk-free bond. Define a **trading strategy** as a **predictable** vector process $\hat{\phi} = (\hat{\phi}(t), t = 1, \ldots, T)$ with component processes $\phi^{(0)}, \phi^{(1)}, \ldots$, and $\phi^{(m)}$. Here predictable means that $\hat{\phi}$ is $\mathcal{F}_{t-1}$ measurable. Interpret $\phi^{(j)}(t)$ as the quantity of asset j held by the investor between times $t-1$ and t.

Then $\hat{\phi}(t) \cdot \hat{S}(t-1)$ is the value of the investor's portfolio at time $t-1$, and $\hat{\phi}(t) \cdot \hat{S}(t)$ is the value just as time t prices are observed, but before rebalancing the portfolio. Define a trading strategy to be **self-financing** if $\hat{\phi}(t) \cdot \hat{S}(t) = \hat{\phi}(t+1) \cdot \hat{S}(t)$, for $t = 1, \ldots, T-1$. This means there is no further investment or consumption after time 0. Define a trading strategy to be **admissible** if it is self-financing and the **value process** $\Pi_{\hat{\phi}}$ is a positive process where

$$\Pi_{\hat{\phi}}(t) = \begin{cases} \hat{\phi}(t) \cdot \hat{S}(t) & \text{for } t = 1, \ldots, T \\ \hat{\phi}(1) \cdot \hat{S}(0) & \text{for } t = 0. \end{cases}$$

A general **contingent claim** is a nonnegative random variable ψ. Let Φ denote the set of all admissible trading strategies, and let Ψ denote the set of all contingent claims. A contingent claim ψ is **attainable** if there exists some $\hat{\phi} \in \Phi$ such that $\Pi_{\hat{\phi}}(t) = \psi$ with probability 1.

Define an **arbitrage** as some $\hat{\phi} \in \Phi$ such that $\Pi_{\hat{\phi}}(0) = 0$, and for some $T > 0$, $P(\Pi_{\hat{\phi}}(T) \geq 0) = 1$ and $P(\Pi_{\hat{\phi}}(T) > 0) > 0$. Suppose we have an adapted interest rate process, $R(t)$. Define the **discount process**

$$D(t) = e^{\int_0^t R(s)ds}$$

so that $dD(t) = -R(t)D(t)dt$. A probability measure $\tilde{P}$ is **risk-neutral** if P and $\tilde{P}$ are equivalent, and under $\tilde{P}$, the discounted asset price $D(t)S^{(i)}(t)$ is a martingale for $i = 1, \ldots, m$. Then under $\tilde{P}$, the discounted portfolio value $D(t)\Pi_{\hat{\phi}}(t)$ is also a martingale. Let $\mathfrak{P}$ be the set of all risk-neutral probability measures under our market model. The **First**

Fundamental Theorem of Asset Pricing states that if $\mathfrak{P}$ is non-empty, then no arbitrage exists. We say a market is **complete** if every contingent claim is attainable. The **Second Fundamental Theorem of Asset Pricing** states that for a market model that has a risk-neutral probability measure, the model is complete if and only if that risk-neutral probability measure is unique.

2.4 NUMERICAL VALUATION TECHNIQUES

In this section we present an overview of three popular numerical valuation procedures developed to deal with more general price dynamics and boundary conditions, multivariate valuation problems, and options that allow for early exercise. These procedures provide much of the infrastructure around which many contributions to the field over the past thirty years have been centered.

2.4.1 Finite Differences

Schwartz [37] first applied the finite difference technique to solve option valuation problems for which closed form solutions are unavailable. One especially interesting case he considers is valuation of an American option on a stock which pays discrete dividends. The approach provides a general numerical solution to the valuation problem, as well as an optimal early exercise strategy.

As in the Black-Scholes setting, the price dynamic of the underlying asset is a geometric Brownian motion with constant variance rate. However, we consider valuation of options which pay discrete dividends and also allow for early exercise. The partial differential equation which governs the value of the options is the same as in the Black-Scholes setting, however the boundary conditions change due to the dividend payments and early exercise feature. Specifically, let $U(s,t)$ represent the value of an American call option on an asset whose value is s, which expires at time T. The partial differential equation governing the

option value is

$$U_t = tU - rsU_s - \frac{1}{2}\sigma^2 s^2 U_{ss}, \tag{2.22}$$

subject to the boundary conditions:

1. $U(s,0) = \max(0, s-K)$
2. $U(0,T) = 0$
3. $U(s,T^+) = \max(0, s-K, U(s-D,T^-))$
4. $\lim_{s\to\infty} U_s(s,T) = 1$

In condition (3), T^+ and T^- are the instants just before and just after the stock pays the discrete dividend, D. This condition reflects the fact that the asset value drops by D when the dividend is paid, and indicates that it is optimal to exercise the option just before the dividend is paid whenever the cum-dividend exercise value $s-K$ is greater than the ex-dividend option value $U(s-D,T^-)$. Equation 2.22 subject to (1), (2), (3), and (4) has no closed form solution, but can be solved numerically by approximating partial derivatives with finite differences.

We can approximate the derivative of U with respect to s at the point (s,t) by

$$\frac{[U(s+\Delta,t) - U(s,t)] + [U(s,t) - U(s-\Delta,t)]}{2\Delta}$$

The pricing algorithm approximates the partial derivatives at a lattice within the domains of price and time. Consider $n+1$ discrete values

$$s_i = ih, \quad i = 0, \ldots, n$$

in the domain of s, and $m+1$ discrete values

$$t_j = jk, \quad j = 0, \ldots, m$$

in the domain of time. Introduce the notation $U(s_i, t_j) = U_{i,j}$. Then

$$U_s = \frac{[U_{i+1,j} - U_{i-1,j}]}{2h} \tag{2.23}$$

In a similar manner, we can approximate U_{ss} and U_t as

$$U_{ss} = \frac{[U_{i+1,j} - 2U_{i,j} + U_{i-1,j}]}{h^2} \tag{2.24}$$

$$U_t = \frac{[U_{i,j} - U_{i,j-1}]}{k} \tag{2.25}$$

Substituting 2.23, 2.24, and 2.25 into 2.22 yields

$$\frac{[U_{i,j} - U_{i,j-1}]}{k} = rU_{i,j} - \frac{ri[U_{i+1,j} - U_{i-1,j}]}{2} - \frac{\sigma^2 i^2 [U_{i+1,j} - 2U_{i,j} + U_{i-1,j}]}{2}. \tag{2.26}$$

Rearranging terms yields the system of $n-1$ equations in $n+1$ unknowns:

$$a_{1,i}U_{i-1,j} + a_{2,i}U_{i,j} + a_{3,i}U_{i+1,j}, \text{ for } i = 1, \ldots, (n-1) \text{ and } j = 1, \ldots, m, \tag{2.27}$$

where $a_{1,i} = \frac{1}{2}rki - \frac{1}{2}\sigma^2 ki^2$, $a_{2,i} = (1 + rk) + \sigma^2 ki^2$, and $a_{3,i} = -\frac{1}{2}rki - \frac{1}{2}\sigma^2 ki^2$.

We can also write the boundary in finite difference form to yield the additional equations

$$\begin{aligned} U_{0,j} &= 0 \qquad\qquad j = 0, \ldots, m \\ U_{n,j} - U_{n-1,j} &= h \quad j = 0, \ldots, m \end{aligned} \tag{2.28}$$

Equations 2.27 and 2.28 provide a solution for $U_{i,j}$ in terms of $U_{i,j-1}$. Since $U_{i,0}$ is known, the entire set of $U_{i,j}$ can be generated by iterating this procedure, and any desired degree of accuracy can be obtained by choosing h and k sufficiently small (at the cost of increased computational time).

Hull and White [22] suggest a modification to the algorithm which ensures that it converges to the true values of the underlying partial differential equation, and also extend the algorithm to efficiently deal with the case where the value of the option depends on several state variables.

2.4.2 Binomial Pricing Models

In this section we examine the binomial asset pricing model originally presented by Cox, Ross, and Rubenstein in [12]. Although the approach is not relevant to the rest of the proposal, we include the presentation for two reasons: binomial modeling is widely used in practice, and the technique depicts the intuition behind risk-neutral asset pricing. We confine our attention to the single period model for the sake of conciseness.

Consider a market model consisting of a single stock worth s at time 0. At the end of the period, the return on the stock is $r_u - 1$ with probability p or $r_d - 1$ with probability $1 - p$, where $r_u > 1 + r > r_d$. This simple price dynamic can be represented as:

$$s \left\langle \begin{array}{ll} r_u s & \text{with probability } p \\ r_d s & \text{with probability } 1-p \end{array} \right.$$

Let U represent the current value of a call option on the stock which expires at the end of the period, having strike price K. A rational exercise decision at expiration implies that $U(r_u s, T) = \max(0, r_u s - K)$ and $U(r_d s, T) = \max(0, r_d s - K)$.

Suppose we wish to form a portfolio consisting of ϕ shares of the stock and b invested in a risk-free bond. The value process of this portfolio will be

$$\phi s + b \left\langle \begin{array}{ll} r_u \phi s + rb & \text{with probability } p \\ r_d \phi s + rb & \text{with probability } 1-p \end{array} \right.$$

Selecting ϕ and b to equate the end-of-period values of the portfolio to the value of the call for each outcome requires that $r_u \phi s + rb = U(r_u s, T)$ and $r_d \phi s + rb = U(r_d s, T)$. Solving these two equations,

$$\phi = \frac{U(r_u s, T) - U(r_d s, T)}{(r_u - r_d)s} \quad \text{and} \quad b = \frac{r_u U(r_d s, T) - r_d U(r_u s, T)}{(r_u - r_d)r}.$$

Define

$$\tilde{p} = \frac{r - r_d}{r_u - r_d} \quad \text{and} \quad 1 - \tilde{p} = \frac{r_u - r}{r_u - r_d}.$$

It is straightforward to show that if this market model is to exclude arbitrage opportunities, then the value of U at time 0 must be

$$U = \frac{\tilde{p}U(r_u s, T) + (1 - \tilde{p})U(r_d s, T)}{r}.$$

This again is nothing more than the discounted expected value of U at the end of the period under the risk-neutral probability $\tilde{p}$. Somewhat tantamount to the fact that the Black Scholes valuation did not depend on μ, this valuation does not depend on p. Bulls and bears alike must agree on the value of the option. This approach is extended by Hull and White [22] to dividend paying stocks and multivariate valuation problems.

2.4.3 Monte Carlo Approaches

Boyle [6] first applied Monte Carlo techniques to option valuation problems. The approach performs well in pricing European options under a large class of price dynamics, however until recently has not been particularly well-suited for American option valuation problems.

Consider the problem of evaluating the integral

$$\int_A g(x)f(x)dx = \overline{g}$$

where $g(x)$ is an arbitrary function, $f(x)$ is a probability density function, and A is the range of integration. We can obtain a Monte Carlo estimate of $\overline{g}$ by generating a random sample of size n from $f(x)$, and then calculating

$$\hat{g} = \frac{1}{n}\sum_{i=1}^{n} g(x_i). \tag{2.29}$$

Boyle applies this reasoning to the problem of pricing a European call option on a stock which follows a geometric Brownian motion with constant variance rate σ^2. Recall from 2.20 that the price of the option can be written as its discounted expected payoff under the risk-neutral probability measure $\tilde{P}$:

$$u(s, 0) = e^{-rT}\tilde{E}[h(S(T))], \tag{2.30}$$

where $h(s) = \max(s - K, 0)$ is the payoff function. The solution to the stochastic differential equation 2.21 is

$$s(T) = s(0)e^{(r-(\frac{\sigma^2}{2}))T+\sigma z\sqrt{T}},$$

which implies that

$$u(s,0) = e^{-rT}\int_{-\infty}^{\infty} h[s(0)e^{(r-(\frac{\sigma^2}{2}))T+\sigma z\sqrt{T}}]\frac{1}{\sqrt{2\pi}}e^{\frac{-z^2}{2}}dw,$$

where z is distributed standard normally. To obtain an estimate of u of the form 2.30, we simply need to generate a random sample of size n from the standard normal distribution, and then calculate

$$\hat{u}(s,0) = \frac{e^{-rT}}{n}\sum_{i=1}^{n} h[s(0)e^{(r-(\frac{\sigma^2}{2}))T+\sigma z_i\sqrt{T}}].$$

One drawback associated with this type of estimation is that the standard error is inversely proportional to $\sqrt{n}$. Boyle [6] goes on to suggest ways of improving the efficiency of the simulation, namely by introducing control variates and antithetic variates.

Johnson and Shanno [23] introduce a Monte Carlo valuation approach under stochastic volatility. Joy, Boyle, and Tan [24] introduce a quasi-Monte Carlo method of valuing European style options. One major focus of this dissertation will be constructing quasi-Monte Carlo approaches to estimate option prices under various stochastic volatility models, which are more efficient than existing Monte Carlo and other quasi-Monte Carlo approaches.

Longstaff and Schwartz [29] present an American style option valuation algorithm which combines exploitation of the knowledge of the optimal exercise rule at expiration as in finite differencing, with the flexibility and versatility of Monte Carlo simulation. Refinements and extensions of their algorithm will be another focus of the remainder of the dissertation.

3.0 EUROPEAN OPTION VALUATION

The celebrated European option valuation formula of Black and Scholes [5] obtains under the assumption of constant volatility of the underlying asset price process. Since their work, much empirical evidence has discredited the validity of this assumption, instead pointing towards the presence of stochastic volatility amidst asset price processes. Heston [21] introduces what has become a widely used stochastic volatility (SV) model, and also derives a closed-form valuation formula for European call options under his model. Bates [3] introduces a more general class of models (SVJ) which also allows for jumps in underlying asset prices, and also derives a valuation formula for call options under the model.

Meanwhile, the values of many types of derivatives traded in today's markets are not readily available via closed-form valuation formulae, and consequently must be estimated. Examples include derivatives on assets that pay discrete dividends, and exotic derivatives whose payoff functions induce boundary conditions in partial differential equations more complicated than those induced by payoffs of simple calls and puts. In these cases, Monte Carlo procedures have become one very popular method of estimating prices and sensitivities. As depicted in the previous chapter, Boyle [6] applied MC techniques to various option valuation problems. A wide array of extensions and generalizations has since evolved. Among others, Joy, Boyle, and Tan [24] apply quasi-Monte Carlo procedures to option pricing problems, and demonstrate that QMC techniques generally result in faster convergence rates than crude MC techniques. An obvious by-product of faster convergence is that a desired level of accuracy can be obtained using a smaller number of sample points. Hence QMC estimates with an acceptable error bound can generally be obtained in less time than it takes to construct crude MC estimates with the same error. Apart from this benefit however, when using an equal number of sample points QMC estimates generally require at least as much

time to construct as their crude MC counterparts.

Broadie and Kaya [8] introduce exact simulation methods for the SV and the SVJ models. These methods lead to the construction of unbiased MC estimates of option prices. They demonstrate the superiority of using the exact simulation method over biased Euler discretization methods when constructing MC estimates of European call option prices, in many cases both in terms of root MSE and CPU time requirements. However, under the exact method, their soft/hardware still requires a somewhat large amount of CPU time to construct the crude MC estimates, in some cases more than an hour.

In this chapter, we present the Heston stochastic volatility model. We present simulation under the model using Euler discretization and also the exact procedure of Broadie and Kaya. We briefly review some common approaches to QMC integration before presenting a QMC procedure that evaluates integral estimates in dramatically less time than is required by crude MC and other QMC methods. We demonstrate the procedure by constructing option price estimates under the SV model. We also extend the approach to the SVJ model and a case where a closed-form valuation formula is available but cumbersome and not commonly used. Finally we present an alternative simulation method and apply it to valuation under the SVCJ model of Duffie et. al. [13].

3.1 THE HESTON STOCHASTIC VOLATILITY MODEL

The risk neutral stock price and variance processes under Heston's stochastic volatility (SV) model can be represented as

$$dS(t) = rS(t)dt + \sqrt{V(t)}S(t)[\rho dW^{(1)}(t) + \sqrt{1-\rho^2}dW^{(2)}(t)]$$

$$dV(t) = \kappa(\theta - V(t))dt + \sigma_v\sqrt{V(t)}dW^{(1)}(t). \tag{3.1}$$

In the first equation, $S(t)$ is the stock price at time t, $\sqrt{V(t)}$ is the volatility, r is the risk free interest rate, $W^{(1)}(t)$ and $W^{(2)}(t)$ are independent Brownian motions, and ρ is the

instantaneous correlation between the stock price and variance processes. In the second equation, θ represents the long-run mean variance, κ is the force of the reversion towards θ, and σ_v is the volatility of the variance process.

Heston [21] derives a closed-form valuation formula for a European call option on an underlying asset that follows the dynamic 3.1. Black and Scholes [5] and Merton [30] use arbitrage arguments to show that the value of any asset $A(s, v, t)$ must satisfy the partial differential equation

$$\frac{1}{2}vs^2\frac{\partial^2 A}{\partial s^2} + \rho\sigma_v vs\frac{\partial^2 A}{\partial s \partial v} + \frac{1}{2}\sigma_v^2 v\frac{\partial^2 A}{\partial v^2} + rs\frac{\partial A}{\partial s}$$

$$+\{\kappa[\theta - v(t)] - \lambda(s, v, t)\}\frac{\partial A}{\partial v} - rA + \frac{\partial A}{\partial t} = 0, \tag{3.2}$$

where $\lambda(s, v, t)$ represents the price of volatility risk. Similar to the Black Scholes formula, consider a solution to the call option valuation formula of the form

$$u(s, v, t) = p_1 s - p_2 K e^{r(T-t)}, \tag{3.3}$$

where K is the strike price and T is the expiration date. Let $x = \ln(s)$ and

$$u_1 = \frac{1}{2}, \quad u_2 = -\frac{1}{2}, \quad a = \kappa\theta, \quad b_1 = \kappa - \rho\sigma_v, \quad b_2 = \kappa.$$

Substituting 3.3 into 3.2 shows that p_1 and p_2 must satisfy the partial differential equations

$$\frac{1}{2}v\frac{\partial^2 p_j}{\partial x^2} + \rho\sigma_v v\frac{\partial^2 p_j}{\partial x \partial v} + \frac{1}{2}\sigma_v v\frac{\partial^2 p_j}{\partial v^2} + (r + u_j v)\frac{\partial p_j}{\partial x}$$

$$+(a_j - b_j v)\frac{\partial p_j}{\partial v} + \frac{\partial p_j}{\partial t} = 0, \quad \text{for } j = 1, 2. \tag{3.4}$$

For the option price to satisfy standard rational boundary conditions, the PDEs 3.4 must satisfy the terminal condition

$$p_j(x, v, T; \ln[K]) = I_{(\ln[K],\infty)}(x).$$

Although p_1 and p_2 are not available in closed form, their characteristic functions, $f_1(x, v, T; \alpha)$, and $f_2(x, v, T; \alpha)$ also satisfy the PDEs 3.4, subject to the terminal condition

$$f_j(x, v, T; \alpha) = e^{i\alpha x}.$$

The characteristic function solution is

$$f_j(x, v, t; \alpha) = e^{B_1(T-t;\alpha)+B_2(T-t;\alpha)v+i\alpha x}, \tag{3.5}$$

where

$$B_1(\tau; \alpha) = r\alpha i\tau + \frac{a}{\sigma_v^2}[(b_j - \rho\sigma_v\alpha i + d)\tau - 2\ln[\frac{1 - ge^{d\tau}}{1 - g}]],$$

$$B_2(\tau; \alpha) = \frac{b_j - \rho\sigma_v\alpha i + d}{\sigma_v^2}[\frac{1 - e^{d\tau}}{1 - ge^{d\tau}}],$$

$$g = \frac{b_j - \rho\sigma_v\alpha i + d}{b_j - \rho\sigma_v\alpha i - d}, \quad \text{and} \quad d = \sqrt{(\rho\sigma_v\alpha i - b_j)^2 - \sigma_v^2(2u_j\alpha i - \alpha^2)}.$$

One can obtain p_1 and p_2 by inverting the characteristic functions[1]:

$$p_j(x, v, T; \ln[K]) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \text{Re}[\frac{e^{-i\alpha \ln[K]} f_j(x, v, T; \alpha)}{i\alpha}]d\alpha. \tag{3.6}$$

Equations 3.3, 3.5, and 3.6 comprise the European call option valuation formula. The integral in 3.6 can be evaluated using numerical techniques such as the trapezoidal rule to obtain any desired level of precision.

[1]It should be noted that when calculating 3.6, care must be taken to ensure that the calculated characteristic function is continuous. Complex numbers $z = x + yi$ have no unique representation when written as $z = re^{i\theta}$. Most software packages choose the representation satisfying $-\pi < \theta < \pi$, which may cause discontinuities in f_j. In this setting, the issue can be rectified by simply rotating segments of f_j by the appropriate angle. See Appendix C for the option price calculator code.

3.2 SIMULATING ASSET PATHS

In spite of the availability of a closed form valuation formula, it is of practical worth to be able to simulate paths under the dynamic 3.1. We can use simulated paths to construct Monte Carlo estimates of option prices for which closed form solutions as above are unavailable. This is the case with options on discrete dividend paying stocks, and options with various exotic payoff features. We now examine two methods of simulating the stock price and variance processes.

3.2.1 Euler Discretization

One method of approximating paths of the stock and variance processes over a finite time horizon T is Euler discretization. To do so, partition the interval $(0, T)$ into L^* equal segments of length Δt, $0 = t_0 < t_1 < \cdots < t_{L^*} = T$, so that $t_i = iT/L^*$ for $i = 0, \ldots, L^*$. The discretization for the stock price process is

$$S(t_i) = S(t_{i-1}) + rS(t_{i-1})\Delta t + \sqrt{V(t_{i-1})}S(t_{i-1})[\rho\Delta W^{(1)}(t_i) + \sqrt{1-\rho^2}\Delta W^{(2)}(t_i)],$$

and for the variance process is

$$V(t_i) = V(t_{i-1}) + \kappa(\theta - V(t_{i-1}))\Delta t + \sqrt{V(t_{i-1})}\sigma_v\Delta W^{(1)}(t_i),$$

where $\Delta W^{(j)}(t_i) = W^{(j)}(t_i) - W^{(j)}(t_{i-1})$, for $j = 1, 2$.

This procedure can be used to construct Monte Carlo estimates of the value of European style options as described in section 2.4, namely

$$e^{-r(T-t)}E[h(S(T))] \approx \frac{e^{-r(T-t)}}{n}\sum_{k=1}^{n} h(\hat{S}_k(t_{L^*})).$$

Here n is the number of simulated sample paths, $\hat{S}_k(t_{L^*})$ denotes the simulated value of $S(T)$ along the k-th path, and $h(\hat{S}_k(t_{L^*}))$ is the payoff of the option along the k-th path. Although this estimate converges to the true value as n and L^* both approach infinity, it does so quite slowly. In addition to the statistical error of the estimation $\sigma_h/\sqrt{n}$, there is the somewhat

more serious problem of discretization bias associated with the Euler discretization. In lieu of this matter, we shall discuss alternative approaches to European option valuation in the following sections.

3.2.2 Exact Simulation of the Heston Model

Broadie and Kaya [8] provide the following exact method of simulating from the distribution of $S(t)|(S(0), V(0))$ under model 3.1. Conditional on the values of $S(0)$ and $V(0)$, $S(t)$ can be written as

$$S(t) = S(0)\exp[rt - \frac{1}{2}\int_0^t V(s)ds + \rho\int_0^t \sqrt{V(s)}dW^{(1)}(s) + \sqrt{1-\rho^2}\int_0^t \sqrt{V(s)}dW^{(2)}(s)]. \tag{3.7}$$

Conditional on $V(0)$, $V(t)$ can be written as

$$V(t) = V(0) + \kappa\theta t - \kappa\int_0^t V(s)ds + \sigma_v\int_0^t \sqrt{V(s)}dW^{(1)}(s). \tag{3.8}$$

To generate a sample from $S(t)|(S(0), V(0))$, they use the following algorithm:

1. Generate a sample from the distribution of $V(t)|V(0)$.
2. Generate a sample from the distribution of $\int_0^t V(s)ds|(V(t), V(0))$.
3. Recover $\int_0^t \sqrt{V(s)}dW^{(1)}(s)$ from 3.8 using the values of $v(t)$ and $\int_0^t v(s)ds$ from steps 1 and 2.
4. Generate a sample from $\int_0^t \sqrt{V(s)}dW^{(2)}(s)$ given the value of $\int_0^t v(s)ds$ from step 2.

The first step is to generate a value for $V(t)|V(0)$, the distribution of which is a constant multiple of a noncentral chi-squared distribution as demonstrated by Cox et al. [10]:

$$V(t) = \frac{\sigma_v^2(1-e^{-\kappa t})}{4\kappa}\chi'^2_{df}(\lambda) \quad \text{for } t > 0,$$

where

$$df = \frac{4\kappa\theta}{\sigma_v^2} \quad \text{and} \quad \lambda = \frac{4\kappa e^{-\kappa t}}{\sigma_v^2(1-e^{-\kappa t})}V(0) \tag{3.9}$$

are the degrees of freedom and noncentrality parameter of ${\chi'_{df}}^2(\lambda)$.

The next step is to generate a sample from the distribution of $\int_0^t V(s)ds|(V(t), V(0))$. The exact procedure relies on Fourier inversion techniques to invert the characteristic function of the distribution. The Laplace transform

$$E[e^{-a^*}\int_0^t V(s)ds|(V(t), V(0))]$$

stems from the results in Pitman and Yor [35]. The characteristic function is obtained by setting $a^* = ia$:

$$\eta(a) = \frac{\gamma(a)e^{-\frac{1}{2}(\gamma(a)-\kappa)t}(1-e^{-\kappa t})}{\kappa(1-e^{-\gamma(a)t})} \tag{3.10}$$

$$\times \exp[\frac{V(0)+V(t)}{\sigma_v^2}[\frac{\kappa(1+e^{-\kappa t})}{(1-e^{-\kappa t})} - \frac{\gamma(a)(1+e^{-\gamma(a)t})}{(1-e^{-\gamma(a)t})}]]$$

$$\times \frac{I_{.5d-1}[\sqrt{V(0)V(t)}\frac{4\gamma(a)e^{-.5\gamma(a)t}}{\sigma_v^2(1-e^{-\gamma(a)t})}]}{I_{.5d-1}[\sqrt{V(0)V(t)}\frac{4\kappa e^{-.5\kappa t}}{\sigma_v^2(1-e^{-\kappa t})}]},$$

where $\gamma(a) = \sqrt{\kappa^2 - 2\sigma_v^2 ia}$, $d = \frac{4\theta\kappa}{\sigma_v^2}$, and $I_\nu(x)$ is the modified Bessel function of the first kind. They use Fourier inversion to recover the corresponding cumulative distribution function:

$$F(x) = P[\int_0^t V(s)ds|(V(t), V(0)) \leq x] = \frac{2}{\pi}\int_0^\infty \frac{\sin ux}{u}\mathrm{Re}[\eta(u)]du$$

$$\approx \frac{hx}{\pi} + \frac{2}{\pi}\sum_{j=1}^{N_{tr}} \frac{\sin hjx}{j}\mathrm{Re}[\eta(hj)], \tag{3.11}$$

where h is the grid size and N_{tr} is the last value calculated in the trapezoidal rule approximation to the integral. The most computationally demanding part of implementing the trapezoidal rule is calculating η, one factor of which is a modified Bessel function of the first kind with complex argument. This type of Bessel function can be expressed as

$$I_\nu(z) = (\frac{z}{2})^\nu \sum_{j=0}^\infty \frac{(\frac{1}{4}z^2)^j}{j!\Gamma(\nu+j+1)}.$$

Again it should be noted that discontinuities may arise in the characteristic function because z^ν is multi-valued and software packages generally assign the principal value so that $-\pi < \arg(z) \leq \pi$. These discontinuities can no longer be rectified as easily as was the case in the discussion of Heston's closed form solution. To achieve continuity in the characteristic function, it is necessary to use the continuation formula from Abramowitz and Stegun [1] to calculate the value of $I_\nu(z)$ for $\arg(z)$ other than its principal value:

$$I_\nu(ze^{m\pi i}) = e^{m\nu\pi i} I_\nu(z) \quad \{m \text{ an integer}\}.$$

Finally, to obtain a sample value x from the distribution of the integral, one can generate a value U from the Uniform [0,1] distribution, and then use a root finder such as a second order Newton method to determine the value of x such that the approximation in 3.11 is equal to U.

Given the values of $V(t)$, $V(0)$, and $\int_0^t V(s)ds|(V(t), V(0))$ we can use 3.8 to resolve $\int_0^t \sqrt{V(s)}dW^{(1)}(s) = \frac{1}{\sigma_v}[V(t) - V(0) - \kappa\theta t + \kappa \int_0^t V(s)ds]$. Also, since the process $V(t)$ is independent of $W^{(2)}(t)$, the distribution of $\int_0^t \sqrt{V(s)}dW^{(2)}(s)$ given the path from $V(0)$ to $V(t)$ is normal with mean 0 and variance $\int_0^t V(s)ds$. These results yield that $\log S(t) \sim N(\mu(0,t), \sigma^2(0,t))$, where

$$\mu(0,t) = \log S(0) + [rt - \frac{1}{2}\int_0^t V(s)ds + \rho \int_0^t \sqrt{V(s)}dW^{(1)}(s)] \tag{3.12}$$

$$\sigma^2(0,t) = (1-\rho^2)\int_0^t V(s)ds.$$

Lastly, we can generate a value from the distribution of $S(t)|(S(0), V(0))$ by generating a standard normal random variable Z, and setting $S(t) = e^{\mu(0,t)+\sigma^2(0,t)Z}$. We can use the simulated values of $S(t)$ to construct Monte Carlo estimates of European options using the same method as when the sample was generated using Euler discretization.

As mentioned above, a shortcoming of this procedure is that calculating the characteristic function is relatively difficult and time consuming. The calculation at any point in the domain of η involves determining the values of two modified Bessel functions of the first kind, one of which has a complex argument. In order to obtain a desired accuracy, the number of terms

in the sum in 3.11 may be as large as a few hundred. The root finder generally takes 4 to 5 iterations to converge. Collectively these issues result in the usage of a large amount of computation time. It depends to a large extent on N_{tr} and h which themselves depend on the rest of the model parameters, but step 2 generally consumes well over 99% of the total CPU time required to generate a sample from the distribution of $S(t)|(S(0), V(0))$. It is our goal to make better use of this time spent by introducing a quasi-Monte Carlo procedure to construct option price estimates more efficiently than crude Monte Carlo methods.

3.3 QUASI-MONTE CARLO TECHNIQUES

One can construct crude MC estimates of option prices using the values of $S(t)|(S(0), V(0))$ generated in the previous section. For example, define Y_1 to be a random variable that follows the same distribution as $V(T)|V(0)$, Y_2 to be a random variable that follows the same distribution as $\int_0^T V(s)ds|(V(T), V(0))$, and Y_3 to be a random variable that follows the same distribution as $\int_0^T \sqrt{V(s)}dW^{(2)}(s)|\int_0^T V(s)ds$. Let F_{Y_1}, F_{Y_2}, and F_{Y_3} represent the respective distribution functions. Then under model 3.1, the value of a European call option on a stock worth s and whose underlying volatility is $\sqrt{v}$ at time 0 is

$$u(s, v, 0; K) = e^{-rT}\int_{-\infty}^{\infty} h(S(T))dF_{S(T)}(y)$$

$$= e^{-rT}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(S(T))dF_{Y_1}(y_1)dF_{Y_2}(y_2)dF_{Y_3}(y_3) \tag{3.13}$$

where $h(s) = \max(s-K, 0)$, K is the strike price of the option, and T is the expiration date. A crude MC estimate of the option value is given by

$$\frac{e^{-rT}}{n}\sum_{i=1}^{n} h(s_i(T))$$

$$= \frac{e^{-rT}}{n}\sum_{i=1}^{n} h(s(0)\exp[rT - \frac{1}{2}y_{2,i} + \rho\int_0^T \sqrt{v(s)}dW^{(1)}(s) + \sqrt{1-\rho^2}y_{3,i}]), \tag{3.14}$$

where $y_{2,i} \sim Y_2$, and $y_{3,i} \sim Y_3$, for $i = 1, ..., n$. This estimate is unbiased and converges at a rate $O(n^{-1/2})$.

Meanwhile, instead of pseudo-random numbers, QMC methods utilize deterministic sequences of numbers, and generally result in faster convergence than crude MC estimates. Discrepancy is a measure of how well a sequence of numbers fills the unit hypercube. Before considering QMC techniques for option pricing amidst stochastic volatility, we present a brief overview of several types of low-discrepancy sequences often used in QMC estimation.

3.3.1 Faure Sequences

Faure sequences are a common type of low-discrepancy sequences used in QMC integration, especially beneficial in multi-dimensional settings. See Glasserman [15] for a precise definition of discrepancy and a more detailed account of Faure numbers than we provide here.

We begin by showing how to construct a one-dimensional Faure sequence. Let p be a prime number. Then any integer m has a unique expansion in base p. We can generate a quasi-random number in the unit interval [0,1] by mapping the base p expansion of m to a point in [0,1] using the radical inverse function, ϑ_p. In general, m can be expressed in terms of base p as

$$m = \sum_{j=0}^{l} a_{j,p}^{(1)}(m) p^j.$$

and the corresponding quasi-random number is

$$\vartheta_p^{(1)}(m) = \sum_{j=0}^{l} a_{j,p}^{(1)}(m) p^{-j-1},$$

where l is the smallest integer necessary to generate the base p expansion of m. For example, with $p = 3$ and $m = 7$, we have $7 = 2(3^1) + 1(3^0)$. The radical inverse function ϑ_3 "reflects 7 about the decimal point", yielding $\vartheta_3^{(1)}(7) = 1(3^{-1}) + 2(3^{-2}) = \frac{5}{9}$. Numbers in the sequence $\vartheta_3^{(1)}(i)$ for $i = 1, ..., 9$ are $(\frac{9}{27}, \frac{18}{27}, \frac{3}{27}, \frac{12}{27}, \frac{21}{27}, \frac{6}{27}, \frac{15}{27}, \frac{24}{27}, \frac{1}{27})$. Notice how numbers added to the sequence fill the gaps in the existing sequence, which makes it possible to terminate a simulation once a desired level of accuracy is obtained.

To construct a d-dimensional Faure sequence, coordinates of the first dimension are constructed exactly as above. Coordinates of the remaining dimensions are constructed recursively. Suppose we know all of the $a_{j,p}^{(k-1)}(m)$. Then

$$a_{j,p}^{(k)}(m) = [\sum_{i \geq j}^{l} [\frac{i! a_{j,p}^{(k-1)}(m)}{j!(i-j)!}]] \mod p,$$

$$\vartheta_p^{(k)}(m) = \sum_{j=0}^{l} a_{j,p}^{(k)}(m)^{-j-1}, 2 \leq k \leq d.$$

Note that for multi-dimensional Faure sequences of certain lengths, the coordinates of dimensions $2, ..., d$ are simply permutations of the coordinates of the first dimension. In fact, the i-th coordinates of any two elements of a d-dimensional Faure sequence cannot be the same.

3.3.2 Panel Numbers

Another type of one-dimensional sequence used in QMC integration can be constructed using the n panel method (or midpoint rule). To construct a one-dimensional sequence of panel numbers, first divide the unit interval into n equal subintervals, $(0, \frac{1}{n}), (\frac{1}{n}, \frac{2}{n}), ..., (\frac{n-1}{n}, 1)$, and then take the midpoint of each subinterval. Niederreiter [34] shows that minimum discrepancy over all one-dimensional sequences of length n is obtained using panel numbers. Evaluation of one-dimensional integrals using panel numbers results in convergence at rate $O(n^{-1})$, as opposed to $O(n^{-1/2})$ under crude MC methods. A shortcoming however, is that the sequence of the first $n+1$ panel numbers has no elements in common with the sequence of the first n panel numbers. This makes it implausible to continue a simulation "just until" a desired level of accuracy is obtained. Another shortcoming of panel numbers arises when we consider the multi-dimensional generalization. Letting ${}_n\xi = \left\{\xi^{(1)}(i)\right\}_{i=1}^{n} = \left\{\frac{2i-1}{n}\right\}_{i=1}^{n}$ represent the sequence of the first n one-dimensional panel numbers, we can construct a d-dimensional sequence by crossing ${}_n\xi$ with itself d times, forming a d-dimensional grid in the unit hypercube. Tavella [41] demonstrates that these d-dimensional panel numbers fall victim to the curse of dimensionality, and do not do a good job of filling the unit hypercube. For these and other reasons, QMC sequences of this type are rarely used to estimate

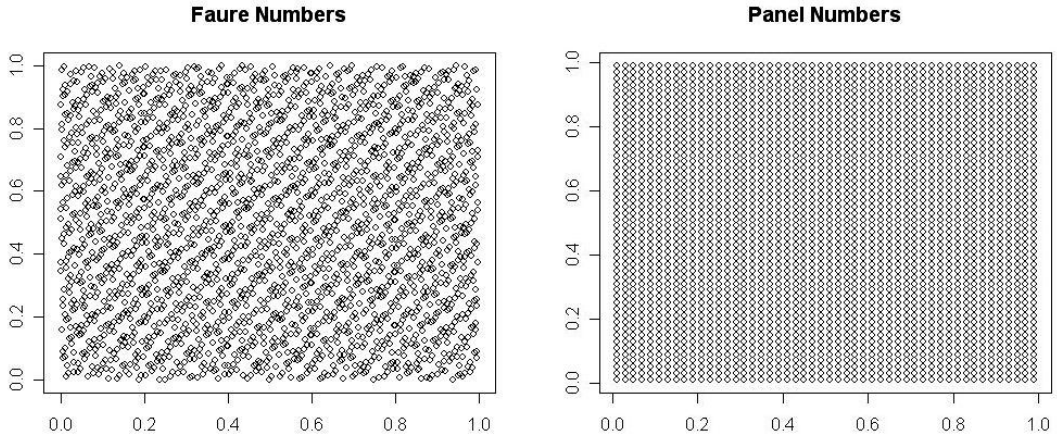

Figure 2: The first 2500 points of each type of sequence.

multi-dimensional integrals. Be that as it may, in the next section we demonstrate the considerable gains in efficiency one can achieve by using multi-dimensional panel sequences in conjunction with simulations that utilize Fourier inversion of a characteristic function. For illustrative purposes, Figure 2 depicts the differences between Faure sequences and panel number sequences in two dimensions. In the figure and all examples throughout the paper, we construct Faure sequences using $p = 5$. Results similar to ours hold for other values of p.

3.4 QMC OPTION PRICE ESTIMATION UNDER THE SV MODEL

In this section we present two QMC estimates of European call option prices under model 3.1. We draw comparisons in terms of root MSE and CPU time requirements between price estimates obtained using the QMC approaches and crude MC.

3.4.1 Option Price Estimates Using Faure Sequences

We see in 3.13 that the value of a European call option under the SV model can be written as a 3-dimensional integral. Let $\left\{\vartheta_p^{(1)}(i), \vartheta_p^{(2)}(i), \vartheta_p^{(3)}(i)\right\}_{i=1}^{n}$ represent the sequence of the

Table 1: Crude MC vs. QMC Faure.

n	*RMSE Crude*	CPU_C	*RMSE Faure*	CPU_F
8,000	0.0839	3.7	2.7×10^{-4}	7.4
64,000	0.0297	29.6	9.8×10^{-4}	59.2
216,000	0.0161	99.9	3.1×10^{-4}	199.9
512,000	0.0105	236.8	7.5×10^{-5}	473.9
1,000,000	0.0075	462.5	6.1×10^{-5}	925.6
8,000,000	0.0027	3701.1	3.4×10^{-5}	7405

Model parameters: $s(0) = 100, v(0) = 0.010201, \kappa = 6.21, \theta = 0.019, \sigma_v = 0.61, \rho = -0.7, r = 3.19\%; K = 100, T = 1$ year, true option price = 6.8061.

first n elements of a 3-dimensional Faure sequence. A QMC estimate of the option price in 3.14 is given by

$$\frac{e^{-rT}}{n} \sum_{i=1}^{n} h(s(0) \exp[rT - \frac{1}{2} F_{Y_2,i}^{-1}(\vartheta_p^{(2)}(i)) + \rho \int_0^T \sqrt{v(s)} dW^{(1)}(s) + \sqrt{1-\rho^2} F_{Y_3,i}^{-1}(\vartheta_p^{(3)}(i))]). \tag{3.15}$$

Note that each iteration throughout the summation leads to different distributions $F_{Y_2,i}^{-1}$ and $F_{Y_3,i}^{-1}$, because the values of $v(t)$ and $\int_0^T v(s)ds$ change with each iteration, hence the addition of the subscript to the notation for the inverse distribution functions.

The estimate in 3.15 converges at a rate $O(n^{-1}(\log n)^3)$. Table 1 compares root MSE and CPU times [2] of option price estimators using crude MC and QMC with Faure sequences. The parameter values are estimated by Duffie, Singleton, and Pan [13] using market option prices for the S&P 500. To be clear, RMSE of the (unbiased) MC procedure is the same as the standard error of the estimate. What we refer to as RMSE of the (deterministic) QMC

[2] To construct all estimates in this paper, we used a desktop PC with a 2.0 GHz Intel Core 2 Duo processor with 800 MHz FSB, 1 GB of RAM, and Windows XP Professional. All programs were written in C and compiled using the Dev C++ compiler.

procedure is simply the error of the integral estimate. QMC Faure clearly outperforms crude MC in terms of RMSE, but the QMC estimates take roughly twice as long to construct as crude MC estimates. This is due to the large amount of time required to determine quantiles of a non-central chi-square distribution relative to the time required to simply generate pseudo-random non-central chi-square observations.

3.4.2 Option Price Estimates Using Panel Numbers

In addition to requiring precise price estimates, practitioners also require quick estimates. The crude MC estimates of the previous section have CPU time requirements ranging from 3.7 seconds for 8,000 paths to 3701.1 seconds for 8,000,000 paths. The QMC Faure method of Section 3.4.1 was very effective in reducing RMSE of the option price estimates, unfortunately the estimates took twice the time to compute as the crude MC estimates. Considering the rate at which market conditions change and pending the level of precision required, this may be a serious issue. In this section we present a QMC procedure which significantly reduces the CPU time required to compute price estimates.

As mentioned above, simulating from the distribution of $\int_0^T V(s)ds|(V(T), V(0))$ consumes well over 99% of the total CPU time required by the overall algorithm. This is because of the time it takes to evaluate the characteristic function and the possibly large number of times it needs to be evaluated in order to obtain an accurate trapezoidal rule approximation to the distribution function in 3.11,

$$F(x) \approx \frac{hx}{\pi} + \frac{2}{\pi} \sum_{j=1}^{N_{tr}} \frac{\sin hjx}{j} \mathrm{Re}[\eta(hj)].$$

Observe that within the approximation which stems from the Fourier inversion, x is not associated with the values of the argument of the characteristic function at which it needs to be evaluated. This suggests that for fixed values of $v(u)$ and $v(t)$, if we wished to determine values of, say $x_1, ..., x_n$ corresponding to numbers in the unit interval $u_1, ..., u_n$, we could do so by evaluating the characteristic function only N_{tr} times rather than $N_{tr}n$ times. In order for such an approach to be plausible, the points in the unit cube we use to construct sums of the form in 3.15 need to be appropriately located. That is to say that many

points having different second coordinates need to have the same first coordinate, and the points also need to be (at least somewhat) nicely spaced throughout the unit cube. As mentioned earlier, Faure sequences lack the characteristic of common coordinates, however this is a feature inherent to the multi-dimensional panel number sequences constructed in the previous section. It remains to be seen whether the fact that the 3-dimensional panel number sequences do not fill the unit cube as uniformly as 3-dimensional Faure sequences will be a hindrance.

Choose an integer n that is a perfect cube, and let ${}_{n^{1/3}}\xi = \left\{\xi^{(1)}(i)\right\}_{i=1}^{n^{1/3}}$ represent the associated one-dimensional sequence of panel numbers. Let ${}_{n^{1/3}}\xi \times_{n^{1/3}} \xi \times_{n^{1/3}} \xi = \left\{\xi^{(1)}(i), \xi^{(2)}(j), \xi^{(3)}(k)\right\}$ for $i, j, k = 1, ..., n^{1/3}$ represent the corresponding 3-dimensional sequence. A QMC estimate of the option price in 3.14 is now given by

$$\frac{e^{-rT}}{n} \sum_{i=1}^{n^{1/3}} \sum_{j=1}^{n^{1/3}} \sum_{k=1}^{n^{1/3}} h(s(0) \exp[rT - \frac{1}{2} F_{Y_{2,i}}^{-1}(\xi^{(2)}(j)) + \rho \int_0^T \sqrt{v(s)} dW^{(1)}(s) + \sqrt{1-\rho^2} F_{Y_{3,i}}^{-1}(\xi^{(3)}(k))]). \tag{3.16}$$

In Figure 3 we provide schematics of the two different QMC procedures used to construct the estimates 3.15 and 3.16. In the first stage in each diagram we are sampling from $V(T)|V(0)$, in the second stage from $\int_0^T V(s)ds|(V(T), V(0))$, and in the final stage from $\int_0^T \sqrt{v(s)} dW^{(2)}(s)| \int_0^T V(s)ds$. At the end of the first stage is where we need to evaluate the characteristic function η in 3.11. Note that under the QMC grid method we are able to consider n values of $S(T)$ by evaluating η $N_{tr} n^{1/3}$ times. Meanwhile, to consider n values of $S(T)$ under the QMC Faure (and crude MC) method, we must evaluate η $N_{tr} n$ times. This fact leads to dramatic reductions in CPU time requirements. Furthermore, determining quantiles of a non-central chi-square distribution is time consuming. Just as was the case with η, QMC grid method requires $n^{1/3}$ quantiles from the non-central chi-square distribution, whereas QMC Faure requires n.

Table 2 provides comparisons of RMSE values of option price estimators under crude MC and each of the QMC procedures. The RMSE values are obtained using $N_{tr} = 200$ and $h = 15$ which, under this set of parameter values, provide four decimal place accuracy for the values in the sample from the distribution of the integral in step 2. Not surprisingly,

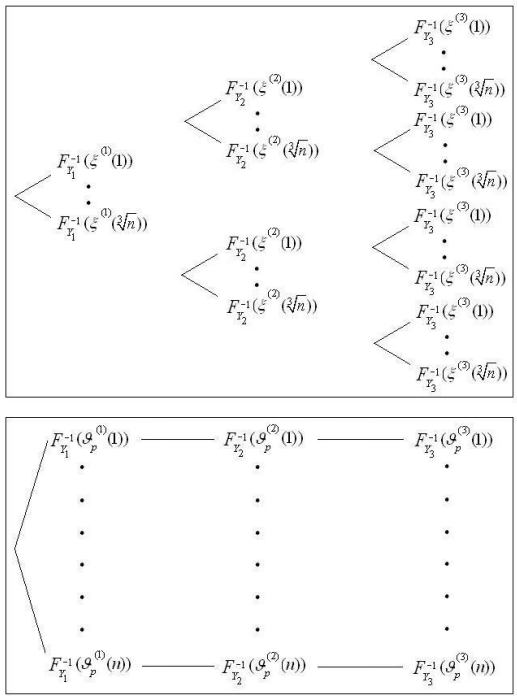

Figure 3: The first diagram illustrates the QMC grid procedure, the second illustrates the QMC Faure procedure.

we see that the QMC grid method fails to perform well strictly in terms of RMSE, in most cases yielding values even greater than the crude MC method. However the benefit of the grid approach becomes evident when comparing CPU time requirements under each method. In Figure 4 we illustrate direct comparisons of CPU time requirements of the three estimators. Table 2 reports the times as ratios under the column headings $\text{CPU}_{C:G}$ and $\text{CPU}_{F:G}$. The values under $\text{CPU}_{C:G}$ are the ratios of CPU time under crude MC to CPU time under the QMC grid method. The values under $\text{CPU}_{F:G}$ are the ratios of CPU time under the QMC Faure method to CPU time under the QMC grid method. These ratios lead to standard efficiency comparisons of the estimators. Suppose two simulation estimators have different root MSE values, RMSE_1 and RMSE_2. Also suppose the CPU times required by the estimates are b_1 and b_2. Then estimator 1 should be preferred over estimator 2 if $\text{Eff}_{2:1} = \text{RMSE}_2 b_2 / \text{RMSE}_1 b_1 > 1$. In terms of this measure of efficiency, the

Table 2: Numerical comparisons under the SV model.

n	*RMSE Crude*	*RMSE Faure*	*RMSE Grid*	$CPU_{C:G}$	$CPU_{F:G}$	$Eff_{C:G}$	$Eff_{F:G}$
8,000	0.0839	2.7×10^{-4}	0.0653	76.8	153.8	98.7	0.64
64,000	0.0297	9.8×10^{-4}	0.0326	172.2	344.6	156.8	10.3
216,000	0.0161	3.1×10^{-4}	0.0259	268.1	536.7	166.7	6.4
512,000	0.0105	7.5×10^{-5}	0.0160	363.6	727.7	238.6	3.4
1,000,000	0.0075	6.1×10^{-5}	0.0127	454.2	909.1	268.2	4.4
8,000,000	0.0027	3.4×10^{-5}	0.0062	899.2	1799.7	391.6	9.9

Model parameters: $s(0) = 100, v(0) = 0.010201, \kappa = 6.21, \theta = 0.019, \sigma_v = 0.61, \rho = -0.7, r = 3.19\%; K = 100, T = 1$ year, true option price = 6.8061.

QMC grid method dramatically outperforms crude MC and also performs quite well against QMC Faure. The disparity encountered when drawing comparisons strictly in terms of RMSE is easily rectified when we incorporate CPU time requirements as part of the method

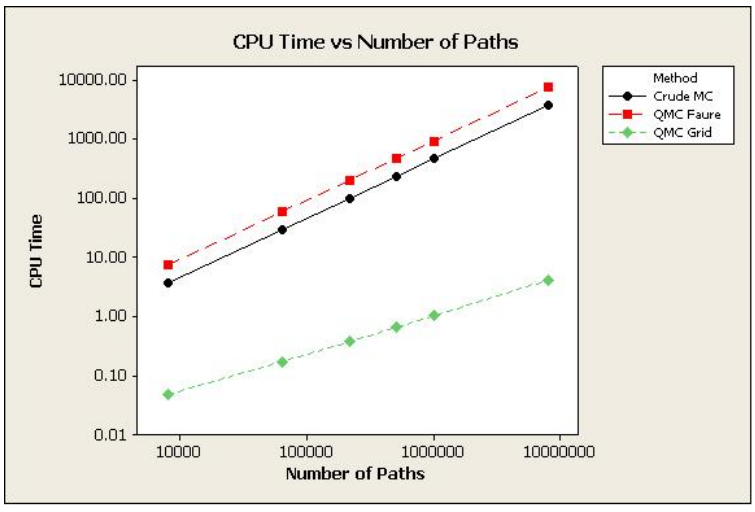

Figure 4: Comparisons of CPU time requirements.

Table 3: Numerical comparisons under the SV model.

n	*RMSE Crude*	*RMSE Faure*	*RMSE Grid*	$CPU_{C:G}$	$CPU_{F:G}$	$Eff_{C:G}$	$Eff_{F:G}$
8,000	0.6848	0.0646	1.3806	51.7	167.0	25.6	7.8
64,000	0.2421	0.0287	0.7685	120.0	387.5	37.8	14.5
216,000	0.1319	0.0080	0.5450	196.6	634.9	47.6	9.3
512,000	0.0859	0.0033	0.4268	261.9	845.9	52.7	6.5
1,000,000	0.0613	0.0007	0.3529	1488.5	4792.7	56.4	2.1
8,000,000	0.0217	0.0003	0.1953	652.9	2108.4	72.5	3.2

Model parameters: $s(0) = 100, v(0) = 0.09, \kappa = 2, \theta = 0.09, \sigma_v = 1, \rho = -0.3, r = 5.0\%; K = 100, T = 5$ years, true option price = 34.9998.

of comparing the three estimators. A similar manner of comparison stems from comparing RMSE values while allocating the same computational budget to each method. In 1 second of CPU time, we can use about 2,160, 1,080, and 1,000,000 sample points under MC, QMC Faure, and QMC grid, respectively. The corresponding RMSE values are 0.1613, 0.0196, and 0.0127. We again note that when using the QMC grid method, we essentially lose the ability to terminate the simulation as soon as we obtain a desired level of precision. Given the fact the QMC methods actually often lead to unreliable error bounds, this matter is relatively inconsequential, especially in lieu of the otherwise considerable gains in efficiency.

In Table 3 we provide numerical results for an artificial and somewhat recalcitrant set of parameter values. All three methods result in slower convergence, with the QMC grid method again performing worst in terms of RMSE. For this set of values, we obtain four decimal place accuracy in the integral sample using $N_{tr} = 80$ and $h = 1$. This explains the differences in CPU time ratios from those under the previous set of parameters. We see again that even amidst very mediocre performance strictly in terms of RMSE, the reduction in CPU time under the QMC grid method leads to substantial gains in efficiency. For the practitioner whose main concern is small RMSE, we note that one can use a billion sample

points under the QMC grid method resulting in RMSE of 0.025 in roughly 60 seconds. The RMSE of the QMC Faure estimate in the second row of Table 3 is 0.0287 and takes roughly 50 seconds to calculate. The efficiency measures of these two estimators are essentially equal.

3.4.3 Multi-Dimensional Panel Numbers in Other Settings

3.4.3.1 The SVJ Model The stochastic volatility with jumps model (SVJ) of Bates [3] generalizes the SV model to allow for jumps in the stock price process. The risk-neutral dynamics of the SVJ model are

$$dS(t) = (r - \lambda\overline{\mu})S(t)dt + \sqrt{V(t)}S(t)[\rho dW^{(1)}(t) + \sqrt{1-\rho^2}dW^{(2)}(t)] + (\zeta^s - 1)S(t)dq(t)$$

$$dV(t) = \kappa(\theta - V(t))dt + \sigma_v\sqrt{V(t)}dW^{(1)}(t). \tag{3.17}$$

In the first equation $q(t)$ is a Poisson process with arrival rate λ, and $(\zeta^s - 1)$ is the percentage change in the stock price whenever a jump occurs. The distribution of ζ^s is lognormal with parameters μ_s and σ_s. The parameters $\overline{\mu}$ and μ_s are related by the expression

$$\mu_s = \log(1 + \overline{\mu}) - \frac{1}{2}\sigma_s^2.$$

As in Broadie and Kaya [8], to simulate from $S(t)|(S(0), V(0))$ under the model in 3.17, the diffusion and the jumps can be treated separately. The exact simulation algorithm is:

1. Momentarily disregarding any jumps, update the risk-neutral drift of the price process as given in (3.17) and simulate $S(t)|(S(0), V(0))$ as in Section 3.2.
2. Generate $N \sim$Poisson(λt) to determine the number of jumps over the time horizon.
3. Generate independent jump magnitudes $\zeta_1^s, ..., \zeta_N^s$, from the lognormal distribution with parameters μ_s and σ_s.
4. Calculate the new final stock price $\tilde{S(t)} = S(t)\prod_{i=1}^{N}\zeta_i^s$.

Table 4: Numerical comparisons under the SVJ model.

n	*RMSE Crude*	*RMSE Faure*	*RMSE Grid*
8,000	0.2495	0.0047	0.2150
64,000	0.0883	0.0026	0.0961
216,000	0.0480	0.0008	0.0627
512,000	0.0312	0.0016	0.0450
1,000,000	0.0223	0.0006	0.0351
8,000,000	0.0079	$7.3\times\ 10^{-5}$	0.0125

Model parameters $s(0) = 100, v(0) = 0.008836, \kappa = 3.99, \theta = 0.014, \sigma_v = 0.27, \rho = -0.79, r = 3.19\%, \lambda = 0.11, \overline{\mu} = -0.12, \sigma_s = 0.15; K = 100, T = 5$ years, true option price = 20.1642.

We compute the QMC grid method estimates by first using the methods of Section 3.4.2 to construct the deterministic unadjusted terminal stock price sample. We then incorporate the jumps in steps 2 and 3 via crude MC. We proceed in this manner for several reasons. It enables us to avoid introducing two more dimensions (the number of jumps and the magnitude of the jumps) into the integral in 3.13. Furthermore, simulated numbers of jumps are small, rendering the QMC grid method inadequate to accurately represent the entire distribution of jump magnitudes.

Again using parameter values determined in Duffie et.al. [13], we present RMSE comparisons in Table 4. Since we have introduced randomness into the QMC grid method, the RMSE values have the usual meaning of $\sqrt{bias^2 + variance}$. To estimate this bias and variance, we first construct the deterministic unadjusted terminal stock price sample and then simulate 10,000 different jump scenarios using crude MC as described above. By doing so we are able to estimate bias and variance to obtain the values in the far right column with reasonable accuracy and efficiency. We note that for more than 125,000 trials the variance contribution to RMSE is essentially zero. We observe familiar overall patterns in the RMSE

Table 5: Valuation of forward start calls.

n	*Crude MC (standard error)*	*QMC Faure*	*QMC Grid*
8,000	6.9516 (0.0092)	6.9544	6.9598
64,000	6.9533 (0.0032)	6.9540	6.9569
125,000	6.9548 (0.0023)	6.9540	6.9564

Model parameters: $s(0) = 100, v(0) = 0.010201, \kappa = 6.21, \theta = 0.019, \sigma_v = 0.61, \rho = -0.7, r = 3.19\%; K' = 1, T = 1$ year, true option price = 6.8061.

values. We also again experience significant reductions in CPU time, yielding overall gains in efficiency similar to those observed in the previous SV examples.

3.4.3.2 Pricing Forward Start Options We conclude this section by applying our technique to a setting in which a closed-form option valuation formula exists but is difficult to evaluate. A forward start option is an option whose strike price is determined at a future date T_1. The payoff of a forward start call option is $(S(T_2) - K'S(T_1))^+$, where T_2 is the expiration date and K' is a parameter that determines the eventual strike price, $K'S(T_1)$. As described in Broadie and Kaya [8], the price of a forward start call can be expressed as

$$u_f(s, v, 0; K') = E[e^{-rT_1} S(T_1) u(1, V(T_1), T_1; K')].$$

Under the SV model, the expression $u(1, V(T_1), T_1; K')$ can be evaluated using the Heston formula. Thus $u_f(s, v, 0; K')$ can readily be estimated using any of the MC or QMC techniques discussed throughout this chapter by simulating a sample from $(S(T_1), V(T_1))|(S(0), V(0))$ and calculating the estimated discounted payoff

$$u_f(s, v, 0; K') = \frac{e^{-rT_1}}{n} \sum_{i=1}^{n} s_i(T_1) u(1, V_i(T_1), T_1; K').$$

As we see in Table 5, all QMC estimates lie comfortably within any bid/ask spread of the unbiased MC estimates one would likely encounter for this option in a market environment.

3.5 VALUATION UNDER AN ALTERNATIVE SIMULATION METHOD

In this chapter, we have discussed two ways of simulating asset paths under Heston's stochastic volatility model. Euler discretization methods introduce computational burdens when a large number of sample paths are necessary, and also yield estimates that are biased because of the discretization. The exact simulation method of Broadie and Kaya eliminates the bias associated with the discretization, but generally requires even more computing time than discretization methods. This is in large part due to the time required by the characteristic function inversion step in the simulation algorithm. Earlier in this chapter, we presented a well-performing quasi-Monte Carlo approach that reduced the impact of the computational time required by that inversion step. As we shall see in later chapters, that approach fails to handle American valuation problems. In this section we introduce a different method that reduces computation time requirements by avoiding the inversion step altogether. In addition, the new method will be suitable for American valuation problems.

The exact simulation method of Broadie and Kaya utilizes the hierarchical sampling structure described in Section 3.2 to generate a value of $(S(t), V(t))|(S(0), V(0))$. Two of the three conditional distributions involved are easy to simulate from, but drawing from $\int_0^t V(s)ds|(V(t), V(0))$ is complicated. An approximate value of this integral can instead be constructed by simulating an entire discretized (yet exact) path of the variance process from time 0 to t using 3.8 and 3.9. To do so, partition the interval from 0 to t into L^* segments of length Δt, $0 = t_0 < t_1 < \cdots < t_{L^*} = t$. Now we can generate $V(t_i)|V(t_{i-1})$ for $i = 1, \ldots, L^*$ which will again follow a constant multiple of a noncentral chi-squared distribution:

$$V(t_i) = \frac{\sigma_v^2(1 - e^{-\kappa(\Delta t)})}{4\kappa} \chi_{df}'^2(\lambda_{i-1})$$

where

$$df = \frac{4\kappa\theta}{\sigma_v^2} \quad \text{and} \quad \lambda_{i-1} = \frac{4\kappa e^{-\kappa(\Delta t)}}{\sigma_v^2(1 - e^{-\kappa(\Delta t)})} V(t_{i-1}) \quad \text{for} \quad i = 1, \ldots, L^*.$$

Since $\int_0^t V(s)ds|(V(t), V(0))$ is an ordinary (Lebesgue) integral, an approximate value of

the integral can be obtained using the trapezoidal rule:

$$\int_{approx} = \frac{\Delta t}{2}(V(t_0) + 2V(t_1) + \cdots + 2V(t_{L^*-1}) + V(t_{L^*})). \tag{3.18}$$

This notion leads to an attempt to generate a sample from $(S(t), V(t))|(S(0), V(0))$ as in Section 3.2.2, but by using $\int_{approx}$ in place of $\int_0^t V(s)ds|(V(t), V(0))$ in steps 2, 3, and 4 of the simulation algorithm. We can use this sample to construct Monte Carlo estimates in the usual manner.

Although we need to sample more values from a non-central chi-square distribution, the approximation should achieve some reductions in computational burden since we are by-passing the inversion procedure. But can it do so without increasing bias beyond a tolerable amount? We address this question by using the approximation to simulate European call option prices under the Heston model.

For the sake of straightforward comparison, we focus on two sets of parameter values utilized earlier in this chapter. Tables 6a and 7a depict RMS error of the exact method for various sample sizes. The column labeled steps in Tables 6b and 7b represents the number of steps per year used to construct the variance process in the integral approximation method. The bias column is estimated by simulating a large number of paths (20 million), and taking the differences between the estimated values and the closed form values. For the parameter values in Table 6, the integral approximation method performs very well, especially when 100 or more steps per year are used in constructing the variance paths. In those cases the estimated bias is less than 0.01, and RMS error is very similar to that of the (unbiased) exact simulation method. Furthermore, the ratio of the CPU time required by the exact method to that required by the integral approximation method using 100 steps per year is more than 20. Table 7 yields similar results in terms of RMS error, although for this set of parameters the integral approximation method using 100 steps is only about 10 times as fast as the exact method. This is due to the increase in time until expiration and also to the fact that the number of terms necessary in the trapezoidal rule approximation in the inversion step of the exact method is smaller than the corresponding value for the first set of parameters.

Table 6: Comparing the exact and integral approximation methods.

a. Exact Method

n	*RMSE*
10,000	0.0760
250,000	0.0152
1,000,000	0.0076

b. Integral Approximation

n	*Steps*	*Bias*	*RMSE*
10,000	10	0.047	0.0899
250,000	10	0.047	0.0494
1,000,000	10	0.047	0.0476
10,000	100	0.010	0.0771
250,000	100	0.010	0.0183
1,000,000	100	0.010	0.0126
10,000	1,000	0.0021	0.0761
250,000	1,000	0.0021	0.0153
1,000,000	1,000	0.0021	0.0079

Model parameters: $s(0) = 100, v(0) = 0.010201, \kappa = 6.21, \theta = 0.019, \sigma_v = 0.61, \rho = -0.7, r = 3.19\%; K = 100, T = 1$ year, true option price = 6.8061.

Before proclaiming victory, we should consider the performance of the method throughout the rest of the parameter space. To address this matter, note that the only difference between sampling under the exact method and under the integral approximation method is the way we obtain the value of the Lebesgue integral of the variance path. An inspection of 3.8 and 3.9 reveals that the value of this Lebesgue integral does not depend on $r, \rho, s(0)$, or K. That is to say the error in 3.18 depends only on the parameters $v(0), \kappa, \theta, \sigma_v$, and T. In

Table 7: Comparing the exact and integral approximation methods.

a. Exact Method

n	*RMSE*
10,000	0.6407
250,000	0.1281
1,000,000	0.0641

b. Integral Approximation

n	*Steps*	*Bias*	*RMSE*
10,000	10	0.1652	0.7328
250,000	10	0.1652	0.2183
1,000,000	10	0.1652	0.1796
10,000	100	0.0266	0.7341
250,000	100	0.0266	0.1489
1,000,000	100	0.0266	0.0782
10,000	1,000	0.0165	0.7310
250,000	1,000	0.0165	0.1443
1,000,000	1,000	0.0165	0.0786

Model parameters: $s(0) = 100, v(0) = 0.09, \kappa = 2, \theta = 0.09, \sigma_v = 1, \rho = -0.3, r = 5.0\%; K = 100, T = 5$ years, true option price = 34.9998.

Figure 5, we start with the parameter values in Table 6 and vary these parameters one at a time throughout their relevant ranges of the parameter space. For each parameter value, we construct the closed-form value and the integral approximation option price estimate using 100,000 sample paths. The lines in the plots represent the difference between the closed form values and the integral approximation values. In most cases, the simulated value is within a nickel of the true value; in no cases does it differ by more than a dime.

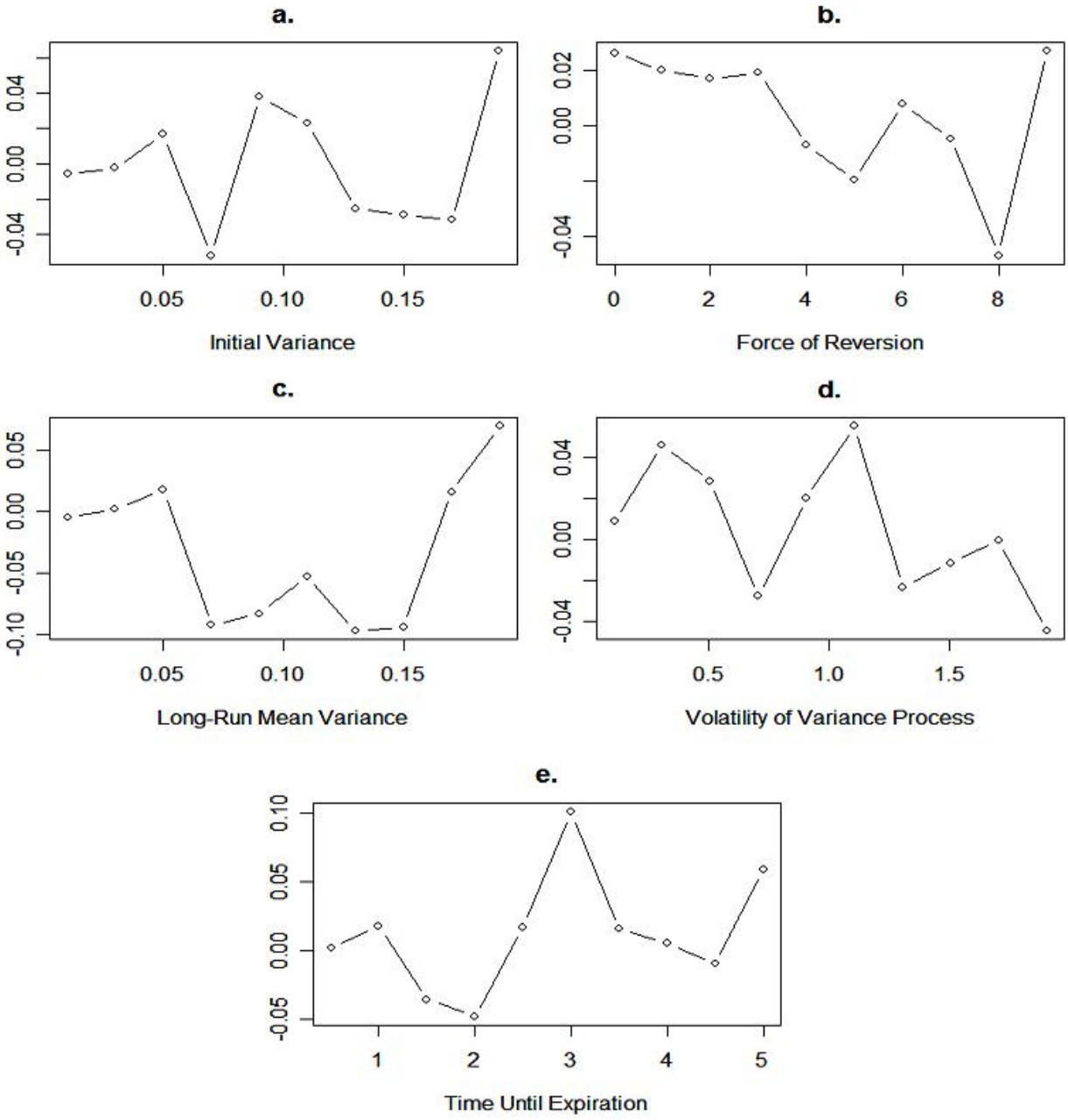

Figure 5: Comparison of closed-form and integral approximation.

We should also consider the purpose of developing such a method in lieu of the fact that a closed-form valuation formula for European call options exists. As we demonstrated at the end of section 3.4, for many exotic options, closed-form valuation formulas are either

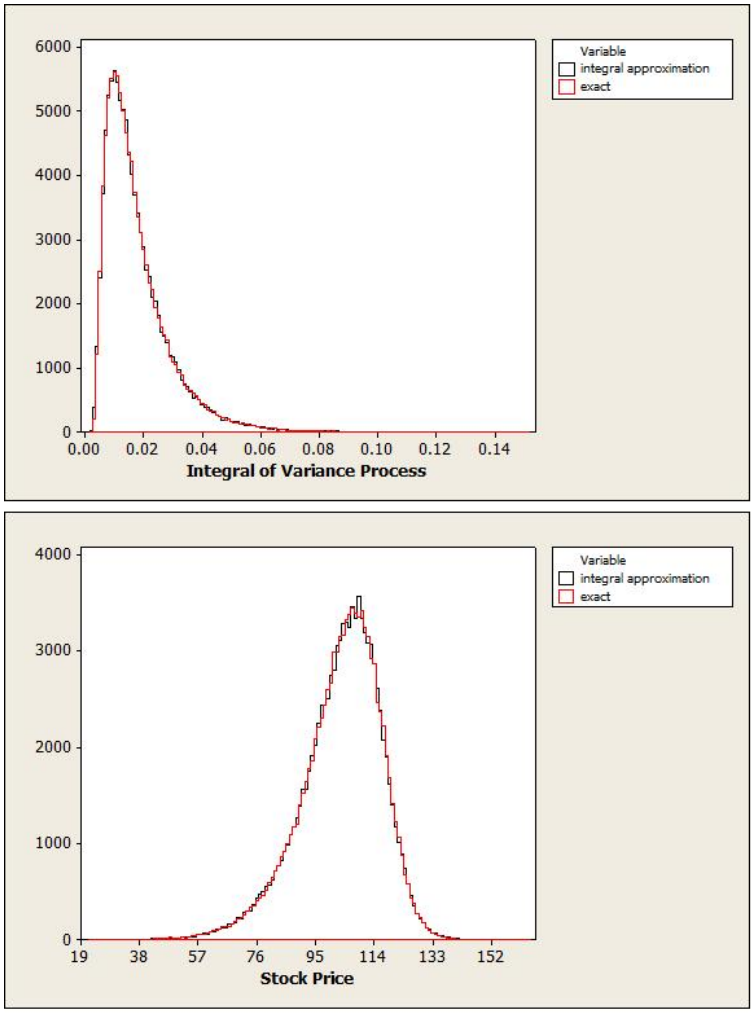

Figure 6: Comparison of exact method and integral approximation.

cumbersome or unavailable. We used the methods of that section to value a forward start call option. The integral approximation method easily extends to valuation under these circumstances. Furthermore, as we shall see in the next chapter, the method of section 3.4 fails when applied to American valuation problems. Meanwhile, the integral approximation method performs quite well.

With that goal in mind, we should also consider the extent to which using the approximate value of the integral of the variance process in place of the true value affects the simulated distributions that depend on the integral. This is important for American option

valuation, because if the distribution of the terminal stock price at the end of one time period varies from the true distribution, then simulating future time points will be unreliable. Figure 6 presents histograms of samples drawn from the distributions of $\int_{t_0}^{t_1} V(s)ds|V(t_0)$ and $S(t_1)|S(t_0)$ for the first set of parameters in Table 6 under both the exact method and the integral approximation. We use 100 steps and $\Delta t = 1$ to construct the variance paths. The histograms appear to be nearly identical, providing strong evidence that any error introduced via the approximation is negligible.

3.6 THE SVCJ MODEL

Before moving to American option valuation in the next chapter, we consider European call valuation under another extension of model 3.1 referred to as the SVCJ model by Duffie, et. al [13]. The model allows for contemporaneous jumps in the stock price and variance processes. This is a phenomenon widely observed in real markets, generally occurring whenever new information becomes available such as a new product launch or news of an acquisition or merger. Simulating under this model to estimate the value of a European option is moving in the direction of American option valuation because we need to simulate values of the stock and variance process at multiple time points rather than simply at option expiration. The risk-neutral dynamics can be written as

$$dS(t) = (r - \lambda\overline{\mu})S(t)dt + \sqrt{V(t)}S(t)[\rho dW^{(1)}(t) + \sqrt{1-\rho^2}dW^{(2)}(t)] + (\zeta^s - 1)S(t)dq(t)$$

$$dV(t) = \kappa(\theta - V(t))dt + \sigma_v\sqrt{V(t)}dW^{(1)}(t) + \zeta^v dq(t).$$

In this dynamic, ζ^v represents the jump size of the variance. The correlation of the simultaneous jumps in stock price and variance is given by ρ_N. ζ^v is assumed to follow an exponential

distribution with mean μ_v. Conditional on ζ^v, ζ^s follows a lognormal distribution with mean $(\mu_s + \rho_N \zeta^v)$ and variance σ_s^2. The parameters μ_s and $\overline{\mu}$ are related by the equation

$$\mu_s = \log[(1 + \overline{\mu})(1 - \rho_N \mu_v)] - \frac{1}{2}\sigma_s^2.$$

We can use the following algorithm to simulate under the SVCJ dynamic using Euler discretization:

1. Simulate a Poisson process with arrival rate λ and determine the time of the next jump; denote this time t_j. If $t_j > T$, then set $t_j = T$.
2. Disregard the jump part, and simulate $V(t_j)$ and $S(t_j)$ using any of the three methods of simulation we have discussed throughout this chapter.
3. If $t_j = T$, set $S(T) = S(t_j)$ and we are finished. Otherwise generate ζ^v by sampling from an exponential distribution with mean μ_v. Update the variance by setting $\tilde{V}(t_j) = V(t_j) + \zeta^v$.
4. Generate ζ^s by sampling from a lognormal distribution with mean $(\mu_s + \rho_N \zeta^v)$ and variance σ_s^2. Update the stock price by setting $\tilde{S}(t_j) = S(t_j)\zeta^s$.
5. Set $s(0) = \tilde{S}(t_j)$, $v(0) = \tilde{V}(t_j)$, $t_0 = t_j$ and go to step 1.

We should note that implementation of Euler discretization of the stock and variance processes is nearly identical to the method described in Section 3.2.1. The discretization for the stock price becomes

$$S(t_i) = S(t_{i-1}) + (r - \lambda\overline{\mu})S(t_{i-1})\Delta t + \sqrt{V(t_{i-1})}S(t_{i-1})[\rho \Delta W^{(1)}(t_i) + \sqrt{1-\rho^2}\Delta W^{(2)}(t_i)],$$

and for the variance process

$$V(t_i) = V(t_{i-1}) + \kappa(\theta - V(t_{i-1}))\Delta t + \sqrt{V(t_{i-1})}\sigma_v \Delta W^{(1)}(t_i),$$

where $\Delta W^{(j)}(t_i) = W^{(j)}(t_i) - W^{(j)}(t_{i-1})$, for $j = 1, 2$.

Table 8 compares valuation of a European call under the SVCJ model using each of the simulation methods (note that the QMC grid methodology of Section 3.4 does not readily apply to this setting for reasons we discuss in the next chapter). The parameter values are

Table 8: Three methods under SVCJ.

a. Exact Method

n	*RMSE*
10,000	0.0720
160,000	0.0184
2,560,000	0.0046

b. Euler discretization

n	*Steps*	*Bias*	*RMSE*
10,000	10/interval	0.0148	0.0744
160,000	40/interval	0.0050	0.0190
2,560,000	160/interval	0.0014	0.0048

c. Integral Approximation

n	*Steps*	*Bias*	*RMSE*
10,000	10/interval	0.0013	0.0735
160,000	10/interval	0.0013	0.0160
2,560,000	10/interval	0.0013	0.0066

Model parameters: $s(0) = 100, v(0) = 0.007569, \kappa = 3.46, \theta = 0.008, \sigma_v = 0.14, \rho = -0.82, \lambda = 0.47, \overline{\mu} = -0.1, \sigma_s = 0.0001, \mu_v = 0.05, \rho_N = -0.38, r = 3.19\%; K = 100, T = 1$ year, true option price = 6.8619.

again taken from Duffie, et.al. [13]. The number of steps used under Euler discretization, as recommended by Broadie and Kaya [8], is equal to $.1 \times \sqrt{n}$ per interval. The number of steps used to construct the variance path in the integral approximation method is 10 per interval, which is the value beyond which we no longer observed greater accuracy. The bias columns are estimated by simulating 20,000,000 paths from which we obtain the estimated

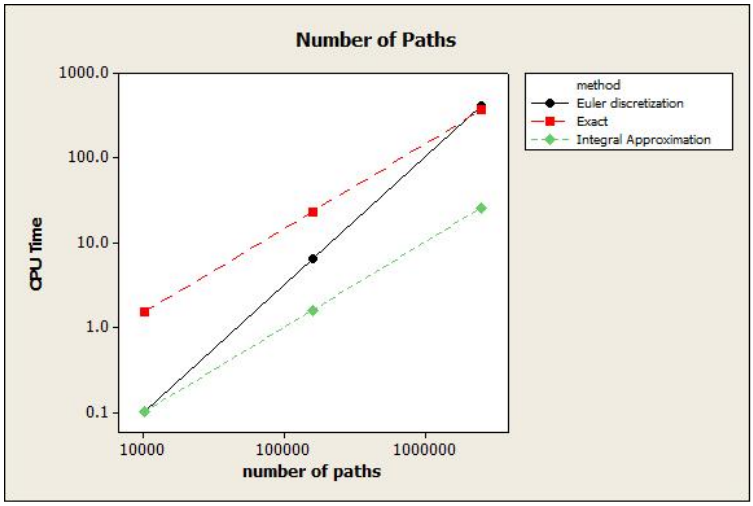

Figure 7: Comparisons of CPU time requirements.

option value, and taking the difference between the estimated and closed-form values.

Although we suppressed comparison between the exact method and Euler discretization earlier in this chapter, the exact method greatly outperforms discretization in those European valuation problems, and this was a major theme in Broadie and Kaya [8]. In the current setting in which we require the simulation of more points along an asset path, we see that the exact method no longer outperforms Euler discretization. The RMSE values of the two methods are nearly identical. Furthermore, when using a practical number of paths the discretization method is faster, as depicted in Figure 7. Meanwhile, the integral approximation method yields similar RMSE values and requires less CPU time than either of the other methods. In this regard, it is more efficient than either of the existing methods. It is also worth noting that the exact simulation method tends to break down numerically whenever the amount of time between succesive jumps is less than about 0.05. This is not an issue under the integral approximation method.

4.0 AMERICAN OPTION VALUATION

In the previous chapter, we presented exact simulation algorithms for jump diffusion processes commonly used to model asset values in the presence of stochastic volatility. We compared crude Monte Carlo European option price estimators with estimators based on quasi-random sequences. One of the QMC approaches dramatically reduced RMSE values, the other dramatically reduced CPU time requirements. We also presented the integral approximation method, which reduced CPU time requirements and yielded RMSE values similar to the exact method.

We now turn our attention to American option valuation. The primary difference between European and American style Monte Carlo option valuation is that American valuation requires the entire simulated path of an asset (suitably discretized), whereas European valuation only requires the terminal value, or at most several values as was the case under the SVCJ model. A popular valuation tool for American style options that will be at the center of our focus is the Least Squares Monte Carlo (LSM) algorithm of Longstaff and Schwartz [29].

The rest of this chapter is organized as follows. In section 4.1, we present an overview of the LSM algorithm. In Section 4.2 we attempt to improve the performance of the LSM algorithm by utilizing quasi-Monte Carlo methods and by replacing the commonly used polynomial basis with a spline basis. In Section 4.3, we discuss constructing paths under the Heston SV dynamic using Euler discretization, the exact simulation algorithm of Broadie and Kaya, and the integral approximation method. In Section 4.4, we apply LSM and compare valuation using sets of paths constructed under the simulation methods discussed in 4.3. We address efficiency of the simulation methods, and sensitivity of the LSM algorithm to the choice of basis functions used in the regressions.

4.1 THE LEAST SQUARES MONTE CARLO ALGORITHM

The American style option valuation algorithm presented by Longstaff and Schwartz in [29] combines the spirit of finite differencing and the versatility of simulation. The algorithm applies to a wide variety of derivatives under a fairly general class of price dynamics. For this reason, along with its simple and efficient implementation, the technique has become increasingly popular among financial engineers.

The algorithm is iterative in nature, and constructs the estimated expected value of continuation of an American option (or other type of derivative) at a given time point, conditional on the option not having already been exercised before that time point. This conditional expectation is estimated using linear regression. Specifically, we first generate a collection of sample paths under the appropriate risk neutral price dynamic. The continuous interval of possible optimal early exercise times in American valuation problems is approximated with a discrete set of time points, and discounted future realized payoffs are regressed on functions of the state variable(s) at each of these time points (using only in-the-money sample paths in the regressions). A complete estimated optimal early exercise strategy obtains under an application of the dynamic programming principle, which implies that one should exercise the option the first time the option is both in-the-money and has an estimated conditional expectation of continuation less than the value associated with immediate exercise. The estimated value of the option is the discounted estimated expected payoff.

4.1.1 A Numerical Example

To help convey the intuition behind the LSM algorithm, consider the following simple numerical example. We wish to price an American style call option on an asset whose current value is 10, and is exercisable at a strike price of 10 at times 1 and 2, where time 2 is the expiration date of the option. For simplicity, suppose the period between each possible exercise date is 1 year and that the annual risk free interest rate is 5%. We wish to apply the algorithm to the collection of simulated sample paths in Table 9 to obtain the stopping rule that maximizes the value of the option along each path. To do so, we need to compute

Table 9: Simulated sample paths.

PATH	$t = 0$	$t = 1$	$t = 2$
1	10	11.02	11.11
2	10	10.66	10.14
3	10	8.99	8.49
4	10	11.96	10.79
5	10	8.31	10.50
6	10	9.44	8.63
7	10	10.08	9.18
8	10	10.67	10.97
9	10	9.24	9.31
10	10	7.55	7.24

several intermediate matrices. The matrix in Table 10 represents the realized cash flows of the option holder from following the optimal exercise strategy at time 2, conditional on the option not being exercised prior to time 2. The optimal strategy at time 2 is simply to exercise the option if and only if it is in-the-money, i.e. if $s(2) > 10$.

Next, if the option is in-the-money at time 1, i.e. along a path such that $s(1) > 10$, the option holder must decide whether to exercise immediately, or to continue without exercising. From the simulated sample path matrix, we see that there are 5 in-the-money paths (1, 2, 4, 7, and 8) at time 1. Let $\tilde{s}(1)$ be a vector of the in-the-money stock prices at time 1 and $\tilde{y}(1)$ the corresponding (discounted) realized future cash flows at time 2 if the option holder chooses not to exercise at time 1 (we don't use out-of-the-money sample paths since a decision whether to exercise isn't relevant). Regressing the discounted cash flows against functions of the state variable gives us an estimate of the conditional expectation of continuation at time 2. As depicted in Table 11, we regress $\tilde{y}(1)$ on a constant and $\tilde{s}_1$, yielding $E[Y|S(1) = s(1)] = -3.6394 + .3873s(1)$. The conditional expectation of continuation is less than the

Table 10: Cash flow matrix at time 2.

PATH	$t = 0$	$t = 1$	$t = 2$
1	-	-	1.11
2	-	-	0.14
3	-	-	0
4	-	-	0.79
5	-	-	0.50
6	-	-	0
7	-	-	0
8	-	-	0.97
9	-	-	0
10	-	-	0

value of immediate exercise whenever $s(1) \geq 10.38$, thus the estimated optimal early exercise rule is to exercise at time 1 in any such case. This suggests that it is optimal to exercise at time 1 along paths 1, 2, 4, and 8, and yields the cash flow matrix in Table 12.

The value of the option is now obtained by simply averaging over all paths the value of the discounted realized cash flows:

$$\frac{\sum_{paths} \text{cash flow}_i \times e^{-.05t_i}}{\#\text{ of paths}} \approx 0.455.$$

4.1.2 Technical Framework and Convergence Results

We assume an underlying probability space $(\Omega, \mathcal{F}, P)$ and time horizon $[0, T]$. Ω is the set of all possible realizations of the stochastic economy on $[0, T]$ and has typical element ω. $\mathcal{F} = \mathcal{F}_T$ is the σ-field of distinguishable events through time T, and P is the probability measure defined on the sets in $\mathcal{F}$. Let $\mathbb{F} = (\mathcal{F}_t; t \in [0, T])$ represent the filtration generated by the associated price dynamic. Consistent with the no-arbitrage paradigm, we assume the

Table 11: Regression at time 1.

PATH	$\tilde{s}(1)$	$\tilde{y}(1)$
1	11.02	$1.11e^{-.05}$
2	10.66	$0.14e^{-.05}$
3	-	-
4	11.96	$0.79e^{-.05}$
5	-	-
6	-	-
7	10.08	0
8	10.67	$0.97e^{-.05}$
9	-	-
10	-	-

existence of a risk neutral probability measure, Q.

The goal is to determine the value of American-style derivative securities which generate random cash flows during $[0, T]$ and have finite-variance random payoffs. The "correct" value is the maximized expected value of the discounted cash flows from the option, where the maximum extends over all stopping times with respect to $\mathbb{F}$. Furthermore, no-arbitrage pricing theory implies that the value of continuation is the expected value of the remaining cash flows with respect to Q when following the optimal stopping rule. Suppose we approximate the continuous interval of possible early exercise times with the discrete set of time points $t_{L-1}, t_{L-2}, \ldots, t_1$. We can then write this expected value as

$$G(\omega; t_l) = E_Q[\sum_{j=l+1}^{L} e^{-\int_{t_l}^{t_j} r(\omega,u)du} CF(\omega, t_j; t_l, T)|\mathcal{F}_{t_l}],$$

where $CF(\omega, s; t, T)$ denotes the path of cash flows generated by the option, conditional on the option not being exercised at or prior to time t, and on the option holder following the optimal strategy at all time points s, $t_l \leq s \leq T$, and $r(\omega, t)$ is the (possibly stochastic)

Table 12: Cash flows realized by following early exercise rule.

PATH	$t = 1$	$t = 2$
1	1.02	-
2	0.66	-
3	-	-
4	1.96	-
5	-	0.50
6	-	-
7	-	-
8	0.67	-
9	-	-
10	-	-

risk-free interest rate.

The LSM approach uses least squares regression to estimate the conditional expectation function at each of the possible early exercise dates. Specifically, at time t_l, assume that the unknown functional form of $G(\omega; t_l)$ can be written as a countable linear combination of $\mathcal{F}_{t_l}$-measurable basis functions. Since we are focusing on derivatives with finite-variance random payoffs, the conditional expectation $G(\omega; t_l)$ will be in the space of square integrable functions under the appropriate measure. This (Hilbert) space will have a countable orthonormal basis, which justifies the assumption that $G(\omega; t_l)$ can be represented as a countable linear combination of elements of this basis.

LSM approximates $G(\omega; t_l)$ by using $M < \infty$ elements from the basis. Denote this approximation as $G_M(\omega; t_l)$. Weak assumptions about the existence of moments imply that the fitted value from the LSM regressions, $\hat{G}_M(\omega; t_l)$, converges in probability and in mean square to $G_M(\omega; t_l)$ as the number of sample paths tends to infinity.

The following result addresses the bias of the LSM algorithm. Let U represent the true

theoretical value of an American option and let n represent the number of simulated sample paths used in the LSM regressions. For any finite choice of M and L, let $LSM(\omega; M, L)$ represent the discounted cash flow associated with sample path ω generated by exercising as soon as the LSM regressions imply that the estimated conditional expectation of continuation is less than the value of immediate exercise. Then

$$U \geq \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} LSM(\omega_i; M, L)$$

with probability one.

Another established convergence result considers a special case in which

$$\frac{1}{n} \sum_{i=1}^{n} LSM(\omega_i; M, L)$$

converges in probability to U. Consider an option on a single underlying asset which follows a Markovian price dynamic, whose price at time t is $S(t)$. Also assume that the option can only be exercised at times t_1 and t_2. If $G(\omega; t_1)$ is absolutely continuous and

$$\int_0^\infty e^{-S} G^2(\omega; t_1) dS < \infty$$

$$\int_0^\infty e^{-S} G_M^2(\omega; t_1) dS < \infty$$

then for any $\epsilon > 0$, there exists a finite M such that

$$\lim_{n\to\infty} P[|U - \frac{1}{n} \sum_{i=1}^{n} LSM(\omega_i; M, L)| > \epsilon] = 0.$$

The important implication of this result is that the number of basis functions used in the regression need not be infinite in order to obtain accurate estimates of U.

Clement, Lamberton, and Protter [9] establish some other LSM convergence results. They demonstrate under fairly general conditions that the entire algorithm converges almost surely, and also determine a rate of convergence and prove that the normalized error is asymptotically Gaussian. Egloff [14] reformulates the optimal stopping problem for Markov processes in discrete time as a generalized statistical learning problem.

4.2 ATTEMPTS TO IMPROVE LSM

In this section, we consider several attempts to improve the performance of the LSM algorithm. First we shall consider a question posed by Longstaff and Schwartz [29], namely whether quasi-Monte Carlo methods result in more efficient estimators. We also consider whether replacing the polynomials commonly used for LSM with a spline basis yields improvements.

4.2.1 QMC for LSM

We shall work under the standard Black Scholes assumption that an underlying asset follows a geometric Brownian motion with constant mean and variance rates. We aim to price American options on the asset using the LSM algorithm, and to determine whether any benefits arise from using quasi-Monte Carlo methodology in place of crude Monte Carlo. For this problem, the reason for implementing quasi-Monte Carlo methodology is to generate paths that are "nicely" spaced throughout the sample space, with the hope that estimates can be obtained with a given accuracy using a smaller sample than required by crude Monte Carlo techniques.

Although the consensus tends to vary, it is generally accepted that quasi-Monte Carlo techniques can yield benefits over crude Monte Carlo when working in a space of up to about 100 dimensions. In the examples we consider, we are estimating the value of American options which expire in one-year. As is common in this setting, we allow for 50 equally spaced early exercise time points, essentially giving the option holder the opportunity to exercise once per week. Given our choice of geometric Brownian motion with constant volatility as the asset price dynamic, the relevant sample space is 50 dimensional.

We could choose to construct the asset paths required to implement LSM using any of the various types of quasi-Monte numbers illustrated by Glasserman in [15], such as Faure sequences, Sobol sequences, and so forth. A problem, however, that arises when attempting to do so for our purpose is that the construction of these numbers yields points for which many or all coordinates lie on the same side of the 45-degree line in a given space. For

example, when using $p = 51$ and $d = 50$ to construct a sequence of Faure numbers, all 50 coordinates of the first number in the sequence are $1/51$. Similarly, all coordinates of the fiftieth number in the sequence are $50/51$. If we were to use these numbers to construct asset paths of geometric Brownian motion over a one-year period in which risk-neutral drift is 0.05, volatility is 0.2, and the initial value is 100, the asset values after one year along those two paths are 5.58 and 1902.95, respectively. Clearly this is inappropriate and will lead to poor estimates.

The question that naturally follows is whether some type of tree similar to that considered in the binomial model in Chapter 2 may be more appropriate. The primary reason trees are inappropriate for our purpose is that in early stages of the tree, many different paths share the same value at a given time point. For example, after the first branch in a binomial tree, all paths have one of only two different possible values. As the LSM algorithm utilizes linear regressions in which the dependent variables are function of state variables, the deficiencies associated with trees for our purpose are apparent.

We now turn to our method of constructing (randomized) QMC paths for LSM. The idea behind our approach is to partially control the volatility of a collection of paths by using a fixed set of quantiles to determine the asset values at each new time point. Recall under a standard geometric Brownian motion that

$$S(t) = s(0)e^{\sigma\sqrt{t}Z+(r-\frac{1}{2}\sigma^2)t}.$$

Suppose we wish to simulate $n = 50$ values at time t. The only random component in the above expression is the standard normal random variable Z. We can create 50 evenly spaced quantiles of the distribution of Z by first dividing the unit interval as we did when constructing midpoint panel numbers in Chapter 3; split the unit interval into 50 equal subintervals and then choose the midpoint of each interval, which in the case of $n = 50$ will yield $1/100, 3/100, \ldots, 99/100$. We then convert to quantiles by applying the standard normal inverse c.d.f. to each of the midpoint numbers. These quantiles can be used to construct 50 quasi-random asset values at time t.

When simulating the change from $t_0 = 0$ to t_1, this method is nothing more than a simple application of the midpoint rule. Meanwhile, after we determine the collection of values at

Table 13: American put valuation under geometric Brownian motion.

Method	*Value*	*s.e.*	*average bias*	*RMSE*
	7.089			
	7.091			
crude MC	7.088	0.0061	0.0078	0.0099
	7.103			
	7.095			
	7.095			
	7.096			
QMC	7.090	0.0035	0.0066	0.0075
	7.092			
	7.099			

Model parameters: $s(0) = 36, \sigma = .4, r = 6\%; K = 40, L = 50, T = 1$ year. Finite difference value = 7.101

time t_1, we need to continue simulating throughout the remainder of the time frame. Our approach is to continue using the midpoint rule and inverse standard normal c.d.f. to obtain quasi-random standard normal variates (in fact the collection is the same at each time point). We need a method, however, of allocating those n numbers to the n paths we are simulating at each time point. We do so simply by generating a random permutation of the integers from 1 to n at each time point and using the corresponding quasi-random standard normal value to determine the change in the value along the path.

To determine the effectiveness of our method, we carried out a simulation study. We implemented LSM to calculate the value of a one-year American put option on a non-dividend paying asset under geometric Brownian motion with $r = 0.06$ and $\sigma = .4$. The value of the asset at the beginning of the period was 36 and the strike price is 40. The results in Moreno and Navas [33] indicate that in this setting results of the LSM algorithm are robust under a large class and number of basis functions. For our basis, we chose polynomials in

the strike price up to the 5th degree. We carried out the valuation 5 times using crude MC and 5 times using our randomized QMC approach, in each case using 100,000 sample paths. Furthermore, as recommended by Broadie and Glasserman [7] and Raymar and Zwecher [36], valuation was conducted out-of-sample, meaning that we first used a simulated collection of 100,000 paths to formulate an estimated optimal stopping rule, and then subsequently simulated another 100,000 paths to which we apply the rule. Results are presented in Table 13. We see that our randomized QMC methodology yields low bias estimates with roughly a 25% reduction in standard error and RMSE when compared to crude MC. This modest reduction stems from the fact that our approach enforces some degree of control over the actual realized volatility of the simulated paths. In fact, we can calculate the average realized volatility over a collection of sample paths using the formula

$$\sigma \approx \frac{1}{n}\sum_{i=1}^{n}\sqrt{\frac{1}{T}\sum_{j=0}^{L-1}(\log\frac{S_i(t_{j+1})}{S_i(t_j)})^2},$$

where n is the number of sample paths in the collection, T is the length of the entire time interval of the simulation, and L is the number of time points used to construct each path. When averaging over the 10 collections of simulated paths under each method, we obtain estimated realized volatility estimates of .4123 and .4021, and the corresponding standard errors are .0092 and .0021. Our randomized QMC method results in realized volatility closer to the true value of 0.4 than the crude MC method.

Another possible approach would be to incorporate stratification into the allocation of the quantiles to the paths. For example, rather than simply generating a random permutation of n quantiles as above and using that permutation to generate asset values, we may stratify the existing collection of paths into a larger half and a smaller half. We can then construct $n/2$ quantiles from the standard normal distribution, and apply two different random permutations of this collection of quantiles to each half of the existing sample paths. We shall reserve exploration of this approach for future work.

4.2.2 LSM Using Splines

In this section, we examine whether any gains in efficiency are achieved by replacing the polynomial basis used in the LSM regressions in the previous section with a spline basis. [1] One reason for doing so is that polynomials do not behave well in the extremes of the range of the state variable, as they tend to be dominated by the component with the largest degree. Another reason is that the boundary conditions imposed by many standard types of payoff functions such as those associated with calls and puts create a 'kink' in an option's value function that may not be well-captured by a polynomial basis.

In Section 2.1, we considered the value of a European call option as a function of time and value of the underlying stock, when that value was assumed to follow a geometric Brownian motion. For the given set of parameters and strike price, Figure 1 depicts the corresponding value function, clearly illustrating the effect the boundary condition associated with the payoff function imposes on the shape of the surface. Note that the kink in the value function is most prevalent when the time until expiration is close to zero and the value of the underlying asset is close to the strike price. Since the asset is assumed to pay no dividends, Theorem 8.5.2 and Corollary 8.5.3 of Shreve [40] imply that the value function of the corresponding American option is the same as that of the European option.

The LSM algorithm, at a given point in time, estimates an option's value as a function of the asset price. It is our goal to demonstrate that a spline basis produces a better fit to the data encountered within the LSM algorithm. In Figure 8, we depict estimates obtained when valuing an American call option using LSM on a non-dividend paying asset which follows a geometric Brownian motion with $r = .05$, and $\sigma = .25$, and for K=15. Corresponding to the collection of simulated paths used in the valuation, the horizontal axis of each plot represents the value of the underlying asset and the vertical axis represents the realized payoff along a particular path. The lines on each plot are fits that represent the conditional expected payoff of the option, given the corresponding value of the state variable. The plots on the left and right column use a polynomial basis and a cubic spline basis, respectively. The plots on the top and bottom row are with 6 months and 1 month remaining until expiration, respectively.

[1] For an overview of spline methodology, see Hastie, et. al. [20]. For a more detailed presentation, see, for example, Green and Silverman [16] and Wahba [42].

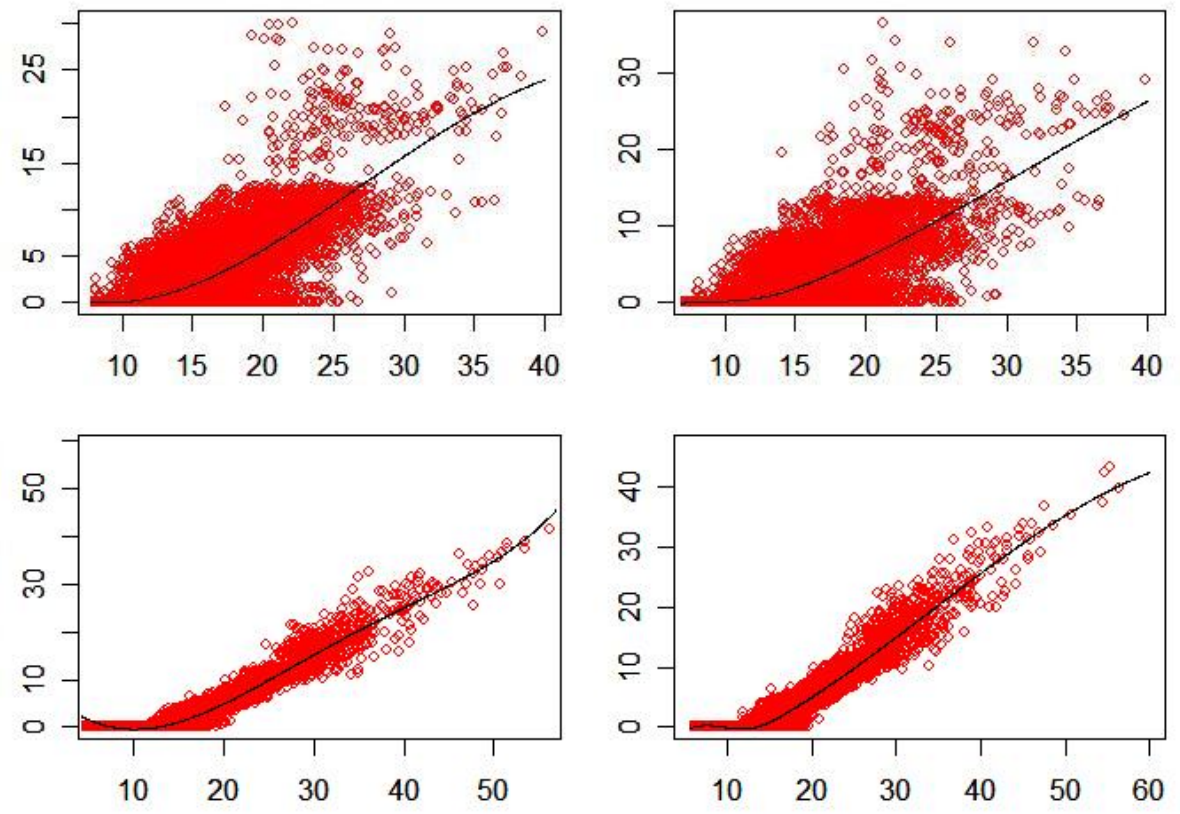

Figure 8: Conditional Expectations Within LSM.

Although there is a high degree of variability in the data around the fitted values, there are benefits in using the spline basis, most notably the behavior of the fit in the extremities of the range of the state variable. For example, when the value of an asset is well above the strike price of an option and there is a short amount of time remaining until expiration, the option's delta is close to 1. The options estimated delta is represented as the slope of the fitted lines, which we can see is clearly closer to 1 for large values of the asset value when using the spline basis.

The above discussion suggests that one may gain some level of efficiency by using splines in conjunction with LSM. For the sake of comparison with the results obtained in the previous section, our investigation begins by again specifying that we are estimating the value of American put options on a single non-dividend paying asset. We also again assume that the

Table 14: American put valuation under geometric Brownian motion.

Basis	*Value*	*s.e.*	*average bias*	*RMSE*
	7.089			
	7.091			
Polynomial	7.088	0.0061	0.0078	0.0099
	7.103			
	7.095			
	7.099			
	7.094			
Spline	7.092	0.0031	0.0072	0.0078
	7.093			
	7.091			

Model parameters: $s(0) = 36, \sigma = .4, r = 6\%; K = 40, L = 50, T = 1$ year. Finite difference value = 7.101

underlying asset follows a geometric Brownian motion with $s(0) = 36, \sigma = .4, r = 6\%$, and that $K = 40$. We shall draw comparisons between valuation carried out by implementing LSM using the same basis that was used in the previous section and the analogous valuation conducted while implementing LSM using a spline basis. To obtain the spline fit, we use a cubic spline with one knot placed at the strike price of the option. This placement makes sense since the kink in the payoff function is located at or near K. The corresponding collection of basis functions is

$$B = \{1, S(t), S^2(t), S^3(t), (S(t) - K)^3\}.$$

Table 14 shows the numerical results when valuing an American put option that expires in 1 year. We again conducted valuation by first simulating 100,000 paths used to construct the stopping rule, and then applying the rule to a different collection of 100,000 paths, repeating this procedure 5 times to obtain estimates of standard error and RMSE. We see a

small reduction in RMSE of about 20% when compared with valuation under the polynomial basis.

Although we shall not do so, one could easily combine the approaches of this section and the last to conduct valuation under a geometric Brownian motion dynamic. Although we were able to obtain modest reductions in RMSE, one drawback of our methodology is that it does not readily extend to more complicated settings involving stochastic volatility and valuation on options on multiple underlying assets, for example. For these reasons, all valuation conducted in the remainder of the dissertation is carried out using standard Monte Carlo construction of asset paths and polynomial bases.

4.3 CONSTRUCTING PATHS UNDER THE SV MODEL

Broadie and Kaya [8] demonstrated the dominance of their exact simulation algorithm over Euler discretization methods, but only in terms of pricing European style options. When used in conjunction with the LSM valuation algorithm of Longstaff and Schwartz [29], both of these simulation methods can be used to price American options. To apply the LSM algorithm under stochastic volatility models, we require the values of the price, and in most cases, the variance processes at the discrete set of time points used to approximate the continuous interval during which early exercise is allowed. The number of time points used in Euler discretization is usually large relative to the number of early exercise points used in the LSM algorithm. In fact, we can easily make the latter a subset of the former, and hence no more work is required when utilizing Euler discretization to observe the entire paths for American valuation than is required to observe only the terminal values of the paths for European valuation. As was depicted in Section 3.6 where we determined European option price estimates under the SVCJ model, requiring multiple values along a path rather than just the endpoint severely reduces the efficiency of the the exact method of Broadie and Kaya. This phenomenon is only exacerbated when we consider American valuation problems.

In this section, we examine the various simulation methods to construct the price and variance processes under the SV model. We first consider the existing methods of Euler

discretization and the exact simulation algorithm of Broadie and Kaya, before presenting our integral approximation method. In the next section we present valuation comparisons.

4.3.1 Euler Discretization

As mentioned in the previous chapter, one method of approximating paths of the stock and variance processes over a finite time horizon T is Euler discretization. To do so, partition the time horizon into L^* equal segments of length Δt, $0 = t_0 < t_1 < \cdots < t_{L^*} = T$, so that $t_i = iT/L^*$ for $i = 0, \ldots, L^*$. The discretization for the stock price process is

$$S(t_i) = S(t_{i-1}) + rS(t_{i-1})\Delta t + \sqrt{V(t_{i-1})}S(t_{i-1})[\rho \Delta W^{(1)}(t_i) + \sqrt{1-\rho^2}\Delta W^{(2)}(t_i)],$$

and for the variance process is

$$V(t_i) = V(t_{i-1}) + \kappa(\theta - V(t_{i-1}))\Delta t + \sqrt{V(t_{i-1})}\sigma_v \Delta W^{(1)}(t_i),$$

for $i = 1, \ldots, L^*$, where $\Delta W^{(j)}(t_i) = W^{(j)}(t_i) - W^{(j)}(t_{i-1}), j = 1, 2$. The time points $t_0, \ldots, t_{L^*}$ can easily be chosen to contain the time points used in the LSM algorithm.

It is worth noting that the price process alone under the SV model is non-Markovian, but when augmented with the variance process the entire process becomes Markovian. When using the LSM algorithm for non-Markovian dynamics, Longstaff and Schwartz recommend including functions of past realizations of state variables as basis functions. We shall later investigate whether it is better to use LSM including past realizations of the price process, augmentation of the variance process, or both.

4.3.2 Exact Simulation

The methods of Broadie and Kaya [8] from the previous section can be easily extended to construct the price and variance processes at all time points required by the LSM algorithm to value American options. To do so, we can simply construct $(S(t_i), V(t_i)|S(t_{i-1}), V(t_{i-1}))$ for $i = 1, \ldots, L$, yielding a discretized (yet exact) realization of the price and variance processes.

It is again worth mentioning that the inversion procedure in the exact simulation method struggles whenever the time increment is less than about 0.05. This is commonly the case when we wish to simulate paths for American valuation, since it is customary to allow early exercise at one week intervals, resulting in time increments of 0.02. Our integral approximation method does not suffer from this deficiency.

4.3.3 Paths Under the Integral Approximation Method

Extending to the construction of entire paths under the integral approximation method of Section 3.5 is similar to extending paths under the exact method. Begin by generating a value of $V(t_1)|V(t_0)$ by simulating the discretized path of the variance as described in Section 3.5. Next we use the variance path to construct the trapezoidal rule approximation to the Lebesgue integral of the variance process using 3.18. After determining the corresponding value of the Ito integral of the volatility process using 3.8, we can simulate a value from the distribution of the stock price at time t_1 using 3.12.

Given these updated values of the variance and the stock price, the same procedure can be repeated to generate paths of the variance and the stock process at each time $t_i, i = 1, \ldots, L$. Repeating the procedure n times yields a collection of paths that can be used in conjunction with the LSM algorithm to arrive at American option price estimates.

4.3.4 The Shortcoming of QMC Grid

In the last chapter, we utilized quasi-random numbers that formed a grid to construct future values of the asset price and variance processes, given the current values. The point of using the grid was so that many points having different third coordinates had the same first and second coordinates, many points with different second coordinates had the same first coordinate, and so forth. This enabled us to construct a collection of n paths without having to carry out each of the time consuming steps of the exact simulation algorithm n times. Because the integral we were evaluating was three-dimensional, there were three values generated in the simulation algorithm for each time point; our grid approach resulted in carrying out the first step $n^{1/3}$ times, the second step $n^{2/3}$ times, and the final (non-

time consuming) step n times, resulting in a sample of size n from the terminal stock price distribution.

This procedure works well when valuing European options, but we encounter difficulties when considering American option valuation. This stems from the increase in the dimensionality of the problem. Suppose we wish to value an option that expires in one year, and that we allow 50 early exercise points in the implementation of the LSM algorithm. This makes the dimensionality of the problem $3 \times 50 = 150$. If we want to construct a sample of size n from the terminal stock price distribution, proceeding as before we would need to begin by generating $n^{1/150}$ points from the distribution in the first step of the algorithm. Clearly this approach is infeasible.

One could develop a sampling procedure in which only certain branches of the grid are carried forward. For instance, one could choose which paths to carry forward using a stratified random sampling mechanism at each branch in the grid. In lieu of the performance of the integral approximation method which we shall present in the next section, however, we shall not pursue such an approach.

4.4 VALUATION COMPARISONS

In this section, we apply the exact method, Euler discretization, and the integral approximation method to various American option valuation problems. We also address the sensitivity of the LSM algorithm to the choice of basis functions used in the regressions.

4.4.1 Valuation of American Call Options Under the SV Model

We now turn to valuation of American call options on non-dividend paying stocks under the SV model. Our first goal is to draw comparisons between the simulation methods discussed in the previous section. We begin by simulating a collection of asset paths under model 3.1 using each method. We can then apply the LSM algorithm to obtain estimated option values under the different methods of simulation. We are able to get an estimate of the

standard error of the overall simulation simply by carrying out this procedure repeatedly. As mentioned previously, the value of an American call option under 3.1 on a non-dividend paying stock is the same as the value of the corresponding European call option. This fact enables us to compare prices for American options obtained via LSM with the closed form European value. Thus we are able to get an estimate of the overall bias of the procedure, and draw comparisons in terms of RMSE.

We should point out before we proclaim one method of simulation better than the rest, that the overall bias of the procedure is not entirely attributable to simulating the asset and variance paths. Clement et. al. [9] prove that under a Markovian dynamic, valuation under the LSM algorithm converges to the true option price as the number of sample paths, n, the number of basis functions used in the regressions, M, and the size of the set of allowable early exercise times, L, all go to infinity. In our setting, under the SV model the stock price process alone is not a Markov process, but if we augment the stock process with the volatility process, the overall process becomes Markovian. It is in general difficult, however, to formulate the bias analytically in terms of n, M, and L, and hence we are unable to separate the overall estimated bias of the procedure into components stemming from the simulation of the paths and the approximations within the LSM algorithm. Nevertheless, since constructing the paths and implementing LSM are both necessary steps in American option valuation via simulation, the overall estimates of bias and RMSE provide for a relevant means of comparison.

In addition to these comparisons, we shall also address the sensitivity of the LSM algorithm to the choice of basis functions used in the regressions. Longstaff and Schwartz [29] suggest that under a non-Markovian dynamic, one should use functions of past values of state variables in the regressions. In light of being able to obtain the Markov property under model 3.1 as described in the previous paragraph, an interesting question is whether it makes a difference to use basis functions that are:

1. Functions of the current price and volatility
2. Functions of current and past values of price
3. Both

Table 15: American call valuation under the exact simulation method.

basis	*estimate*	*bias*	*s.e*	*RMSE*
	6.878			
	6.980			
B_1	7.071	0.181	0.072	0.195
	6.979			
	7.029			
	6.891			
	6.975			
B_2	7.065	0.180	0.062	0.191
	8.987			
	7.023			
	6.884			
	6.989			
B_3	7.074	0.183	0.069	0.195
	6.979			
	7.018			

Model parameters: $s(0) = 100, v(0) = 0.010201, \kappa = 6.21, \theta = 0.019, \sigma_v = 0.61, \rho = -0.7, r = 3.19\%; K = 100, L = 50, T = 1$ year, true option price $= 6.8061$.

As discussed in Section 4.2, types of functions used in LSM regressions are typically polynomials. As mentioned previously, Moreno and Navas [33] show that when valuing standard American style options under the Black-Scholes model, the algorithm is neither sensitive to the type of polynomials used (e.g. power, Hermite, Legendre, etc.) nor the largest degree used (degrees between 3 and 10 yield similar results). We can use their results to guide our choice of collections of functions to use in our comparison. Throughout the remainder of this chapter, our collections of basis functions corresponding to cases 1-3 above are:

Table 16: American call valuation under Euler discretization.

basis	*estimate*	*bias*	*s.e*	*RMSE*
	6.877			
	6.870			
B_1	6.838	0.0503	0.0191	0.0538
	6.834			
	6.861			
	6.853			
	6.851			
B_2	6.821	0.0334	0.0151	0.0367
	6.825			
	6.846			
	6.861			
	6.860			
B_3	6.820	0.0382	0.0191	0.0427
	6.827			
	6.850			

Model parameters: $s(0) = 100, v(0) = 0.010201, \kappa = 6.21, \theta = 0.019, \sigma_v = 0.61, \rho = -0.7, r = 3.19\%; K = 100, L = 50, T = 1$ year, true option price = 6.8061.

- $B_1 = \{S(t), S^2(t), S^3(t), V(t), V^2(t), V^3(t), S(t)V(t), S^2(t)V(t), S(t)V^2(t)\}$
- $B_2 = \{S(t), S^2(t), S^3(t), S(t-1), S^2(t-1), S^3(t-1), S(t-2), S^2(t-2), S^3(t-2), S(t)S(t-1), S(t-1)S(t-2)\}$
- $B_3 = \{S(t), S^2(t), V(t), V^2(t), S(t-1), S^2(t-1), V(t-1), V^2(t-1), S(t-2), S^2(t-2), V(t-2), V^2(t-2)\}$

We should also note that, as above, valuation is conducted out-of-sample. To do so, we formulate an early exercise rule on one set of paths, and then apply that rule to a different

Table 17: American call valuation under the integral approximation method.

basis	*estimate*	*bias*	*s.e*	*RMSE*
	6.805			
	6.821			
B_1	6.786	-0.0163	0.0251	0.0299
	6.779			
	6.755			
	6.817			
	6.831			
B_2	6.798	-0.0056	0.0246	0.0253
	6.785			
	6.769			
	6.815			
	6.830			
B_3	6.797	-0.0079	0.0265	0.0277
	6.784			
	6.762			

Model parameters: $s(0) = 100, v(0) = 0.010201, \kappa = 6.21, \theta = 0.019, \sigma_v = 0.61, \rho = -0.7, r = 3.19\%; K = 100, L = 50, T = 1$ year, true option price = 6.8061.

collection of paths. To obtain an estimate of the standard error associated with the entire valuation algorithm, our methodology is to generate five collections of 100,000 paths, each of which is used to formulate an early exercise rule. We then apply each rule to one of five different simulated collections of 100,000 paths. In order to compare bases, for a given method we use the same collections of paths across all three collections of basis functions under consideration.

The numerical results for the first set of parameters under consideration are presented in Tables 15, 16, and 17. The first column of each table indicates the basis under consideration.

The second column gives the values obtained from the 5 simulation runs, and the last three columns depict bias, standard error, and RMSE of the estimator. We see that the exact simulation method performs somewhat poorly, having RMSE roughly 5 times larger than the other methods. This is a reflection of the fact that the exact algorithm encounters numerical difficulties whenever the simulation interval is small. Here we are valuing an option that expires in one year, and allow for the customary value of 50 early exercise points to do so, meaning that each simulation interval is equal to 0.02. Furthermore, simulating under the exact method requires more than 10 times the CPU time required by either of the other methods. For these reasons, we recommend avoiding the exact simulation algorithm when valuing American options in conjunction with LSM, and shall focus on the alternative methods throughout the remainder of the chapter.

Turning to those methods in Tables 16 and 17, we see that the RMSE under Euler discretization is roughly twice as large as RMSE under our integral approximation approach. As above, we need the values of the asset and variance at each of the 50 time points when we allow for early exercise.[2] Regarding discretization intervals, we use ten subintervals between each early exercise date to carry out the discretization of the asset and variance processes under Euler discretization. We also use ten subintervals when constructing the variance path between early exercise points under the integral approximation approach. These are the values beyond which we observe no further gains in accuracy. These values greatly influence the amount of CPU time required to simulate the paths, and for this combination the integral approximation method is nearly twice as fast as discretization. This is not surprising since between early exercise points, the discretization method constructs a path and variance process whereas the integral approximation method only constructs the variance path. This fact along with the aforementioned reduction in RMSE under the integral approximation method establishes it as an attractive alternative means of simulation under the Heston model when valuing American derivatives.

Recall that our estimates correspond to formulating a stopping rule on one set of 100,000 paths and then applying that rule to an out-of-sample collection of 100,000 paths. Also recall

[2]In addition to the current asset and variance values, B_2 and B_3 utilize two previous values in formulating the early exercise rule. Hence these values are also used when implementing that rule, which reduces the number of allowable early exercise dates to 50-2=48.

Table 18: American call valuation under Euler discretization.

basis	*estimate*	*bias*	*s.e*	*RMSE*
	13.185			
	13.244			
B_1	13.329	0.178	0.0633	0.189
	13.335			
	13.249			
	13.167			
	13.179			
B_2	13.249	0.134	0.0557	0.145
	13.305			
	13.221			
	13.001			
	13.105			
B_3	13.268	0.073	0.1151	0.136
	13.273			
	13.167			

Model parameters: $s(0) = 100, v(0) = 0.09, \kappa = 2, \theta = 0.09, \sigma_v = 1, \rho = -0.3, r = 5\%; K = 100, L = 50, T = 1$ year, true option price = 13.091.

that we use the same collections of paths across bases. This enables us to draw comparisons across the three different bases discussed above. For instance, in Table 17, the first estimates corresponding to each of the three bases are 6.805, 6.817, and 6.815. We see that in this and every other case the estimates are remarkably robust against the choice of basis. Because B_1 and B_3 contain terms corresponding to the latent variance process, that value must be estimated before being able to implement the stopping rules corresponding to those bases, as discussed in the next chapter. On the other hand, B_2 only utilizes observable values of the asset price.

Table 19: American call valuation under the integral approximation method.

basis	*estimate*	*bias*	*s.e*	*RMSE*
	12.997			
	13.161			
B_1	13.037	-0.0235	0.0623	0.0666
	13.087			
	13.048			
	12.993			
	13.119			
B_2	13.029	-0.0358	0.0476	0.0596
	13.076			
	13.051			
	12.990			
	12.983			
B_3	12.994	-0.0856	0.0233	0.0887
	13.012			
	13.041			

Model parameters: $s(0) = 100, v(0) = 0.09, \kappa = 2, \theta = 0.09, \sigma_v = 1, \rho = -0.3, r = 5\%; K = 100, L = 50, T = 1$ year, true option price = 13.091.

When we shift focus to the other set of parameter values under consideration, the comparisons between Euler discretization and the integral approximation remain for the most part in tact. We see that the RMSE of Euler discretization is roughly 2-3 times larger than that of the integral approximation method. Again, the CPU time required by the integral approximation method is only about half that required by Euler discretization. These factors again establish the integral approximation method as competitive among methods of simulating under the SV model.

Meanwhile, our earlier observation that the LSM algorithm was not sensitive to the

Table 20: American put valuation under Euler discretization.

basis	*estimate*	*average*	*s.e*
	4.033		
	4.015		
B_1	4.039	4.038	0.0165
	4.059		
	4.047		
	3.999		
	3.978		
B_2	4.001	4.004	0.0197
	4.033		
	4.011		
	4.004		
	3.980		
B_3	4.026	4.018	0.0258
	4.040		
	4.039		

Model parameters: $s(0) = 100, v(0) = 0.010201, \kappa = 6.21, \theta = 0.019, \sigma_v = 0.61, \rho = -0.7, r = 3.19\%; K = 100, L = 50, T = 1$ year.

differences in the three bases under consideration is no longer valid. We see that B_1 tends to yield larger estimates than B_2, and B_2 yields larger estimates than B_3. No single basis is dominating the others in terms of bias or RMSE across methods. In lieu of the fact that integral approximation seems to be the preferable method of simulation, B_2 and B_3 yield very similar values of RMSE, so again the fact that B_2 does not involve the latent variance process makes it a reasonable collection of basis functions.

Table 21: American put valuation under the integral approximation method.

basis	*estimate*	*average*	*s.e*
	3.967		
	3.949		
B_1	3.967	3.975	0.0202
	3.998		
	3.993		
	3.930		
	3.940		
B_2	3.934	3.949	0.0193
	3.970		
	3.969		
	3.944		
	3.948		
B_3	3.946	3.963	0.0245
	3.999		
	3.978		

Model parameters: $s(0) = 100, v(0) = 0.010201, \kappa = 6.21, \theta = 0.019, \sigma_v = 0.61, \rho = -0.7, r = 3.19\%; K = 100, L = 50, T = 1$ year.

4.4.2 Valuation of American Put Options Under the SV Model

As mentioned previously, for non-dividend paying stocks, the value of an American call option is the same as the value of a European call option. This is not the case, however, for American put options. In this section we use the LSM algorithm to value American put options under the SV dynamic, and again compare the performance of the simulation methods of Section **??**, and also the performance of different bases utilized within the LSM algorithm.

Table 22: American put valuation under Euler discretization.

basis	*estimate*	*average*	*s.e*
	8.991		
	8.896		
B_1	8.965	8.938	0.0390
	8.930		
	8.910		
	8.650		
	8.597		
B_2	8.648	8.618	0.0305
	8.581		
	8.613		
	8.910		
	8.851		
B_3	8.912	8.876	0.0320
	8.849		
	8.858		

Model parameters: $s(0) = 100, v(0) = 0.09, \kappa = 2, \theta = 0.09, \sigma_v = 1, \rho = -0.3, r = 5\%; K = 100, L = 50, T = 1$ year.

In drawing the comparisons, we are hampered by the fact that we have no closed-form value with which to compare our estimated values, and hence we can not calculate bias or RMSE. For the first set of parameters in Tables 20 and 21, as in the previous section, estimated values tend to be somewhat larger under Euler discretization. Standard errors are very similar under each method of simulation. The trend for different bases is similar to that of the previous section, that being a very small difference in estimated values across bases, with B_1 consistently yielding the largest estimates.

We also see a familiar trend for the other set of parameter values under consideration in

Table 23: American put valuation under the integral approximation method.

basis	*estimate*	*average*	*s.e*
	8.774		
	8.644		
B_1	8.766	8.718	0.0596
	8.665		
	8.739		
	8.536		
	8.434		
B_2	8.553	8.501	0.0546
	8.449		
	8.530		
	8.757		
	8.580		
B_3	8.710	8.663	0.0731
	8.604		
	8.666		

Model parameters: $s(0) = 100, v(0) = 0.09, \kappa = 2, \theta = 0.09, \sigma_v = 1, \rho = -0.3, r = 5\%; K = 100, L = 50, T = 1$ year.

Tables 22 and 23. The Euler discretization again yields estimates larger than those obtained under the integral approximation method. In addition, there is a significant difference in the estimates obtained under the different bases. The fact that estimates differ across bases differs from the results obtained by Moreno and Navas [33] under the geometric Brownian motion dynamic. Furthermore, the fact that Euler discretization yields estimates consistently larger than those obtained under the integral approximation method, for both call and put valuation, allows us to conjecture that any biases introduced under these methods of simulation result in excess realized volatility under Euler discretization.

5.0 OTHER RESULTS, FUTURE WORK, AND CONCLUSIONS

As discussed in the previous chapter, American option valuation consists of maximizing an expected payoff, where the expectation is calculated under a risk-neutral probability measure and the maximum extends over all stopping times with respect to a relevant filtration. In as much as knowledge of an optimal stopping rule under a risk neutral measure determines the value of an option, it gives the investor no guidance on how to behave after he establishes a position in an option. In this chapter we examine the utilization of estimated optimal stopping rules to guide an investor's choices regarding an established option position, and consider the return distributions under various formulations of the rule.

5.1 THIN PLATE SPLINES FOR OPTION VALUATION

To develop the methodology described above, we will need an efficient method of valuing American options. LSM was one method of doing so, but as we will see, our methodology will require values of American options at, in some cases, more than 100,000 different combinations of values of state and time. A simple call to LSM on each of these occasions would lead to extreme inefficiencies.

Our method of constructing what amounts to essentially an option valuation formula as a function of state and time will rely on thin plate splines, which are a straightforward extension of natural cubic splines to the multi-dimensional case. We use standard LSM to obtain option valuation estimates at various combinations of values of state and time. We then treat state and time as independent variables and the valuation estimates as dependent variables, and obtain a conditional expectation as a function of state and time by fitting a

thin plate spline surface to the "data". The height of the surface for given values of state and time serves as an estimated option value.

5.1.1 Valuation of American Call Options Under GBM

For this approach to be feasible, obviously the spline fit must yield reliable valuation estimates. To help determine whether this is in fact the case, we can test the procedure on an American call valuation problem. We shall proceed by first using standard LSM to estimate the value of an American call option on a non-dividend paying asset at a collection of values in the (s, t)-plane. We can then fit a thin plate spline to the resulting set of estimated option values. Finally we can test the reliability of the surface as a valuation estimator by comparing fitted values with true option values and also with LSM estimates and associated error bounds. Note that when fitting the splines within this simple framework, we could replace the LSM estimates with the known closed-form values. Doing so, however, is not possible in the settings we consider later in the chapter. Therefore to determine the performance of the surface as a valuation estimator, we choose to fit the surface to simulated values that vary around the true values, as will necessarily be the case in the following sections.

Figure 9 depicts the surface obtained when fitting thin plate splines to a collection of 200 randomly generated values for asset price, time remaining until expiration, and the corresponding LSM value estimates. The parameters of the geometric Brownian motion are $r = 0.05$ and $\sigma = .25$, and the strike price of the option is $K = 15$. We can draw comparisons between this surface and the known true surface depicted in Figure 1. We see that the surfaces share the same shape, the only exception being near the call option boundary condition.[1] To obtain a more quantitative measure of how well the two surfaces coincide, we randomly chose 50 points within the convex hull of the set of points used to generate the spline fit, and compared the fitted values for those 50 points with the true known values obtained from the closed form Black-Scholes equation 2.7. The values differed, on average, by only \$0.0025, a value well within two standard errors of LSM estimates

[1]LSM can only be applied when we observe the simulated collection of paths for at least two periods, hence all of our points that were used to generate the thin plate spline fit have a time value of at least $2/50 = 0.04$. This fact will not be a hindrance since we actually do not need to use the fitted surface to estimate any values along the payoff boundary since those values are known.

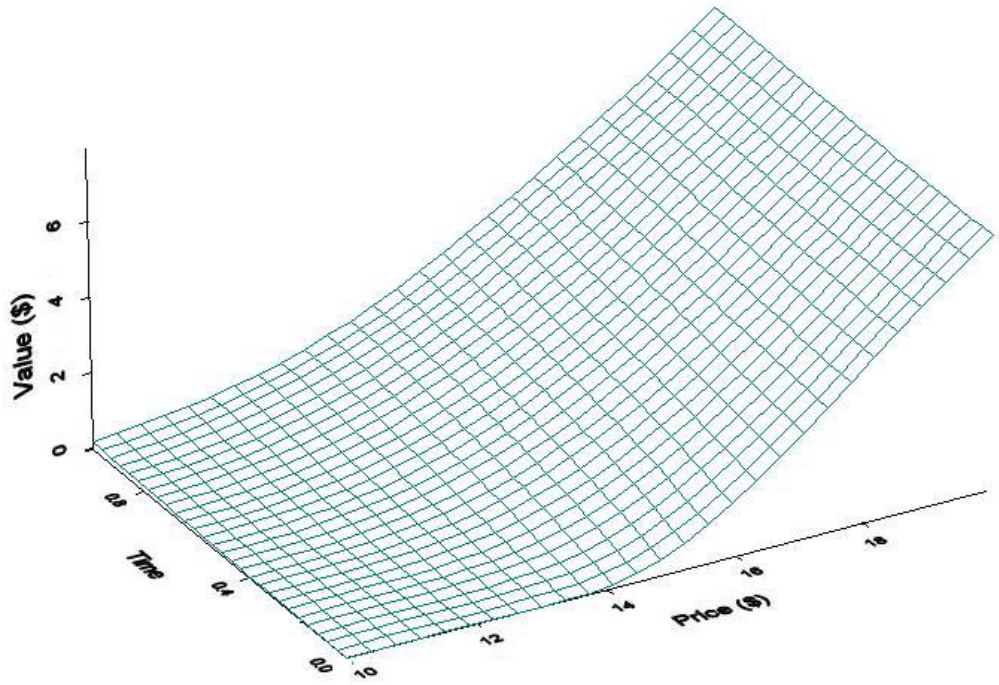

Figure 9: Value surface obtained using thin plate splines under GBM.

obtained for the same points. The maximum deviation from true value to fitted value was \$0.015.

5.1.2 Valuation of American Call Options Under the SV Model

We also wish to consider how well the above approach works in more complex settings. Under SV model 3.1, the American call option valuation formula is a function of underlying asset price, volatility, and time. To obtain an estimated valuation formula, we first utilize LSM to obtain estimated American call option values on a non-dividend paying asset for 1000 different randomly generated values of asset price, volatility, and time. The parameters of the model are assumed to be $\kappa = 6.21$, $\theta = 0.019$, $\sigma_v = 0.61$, $\rho = -0.70$, and $r = 0.0319$, and the strike price of the option is $K = 100$. We then fit a three-dimensional thin plate spline surface to the 1000 points, which is depicted two dimensions at a time in the left column of Figure 10.

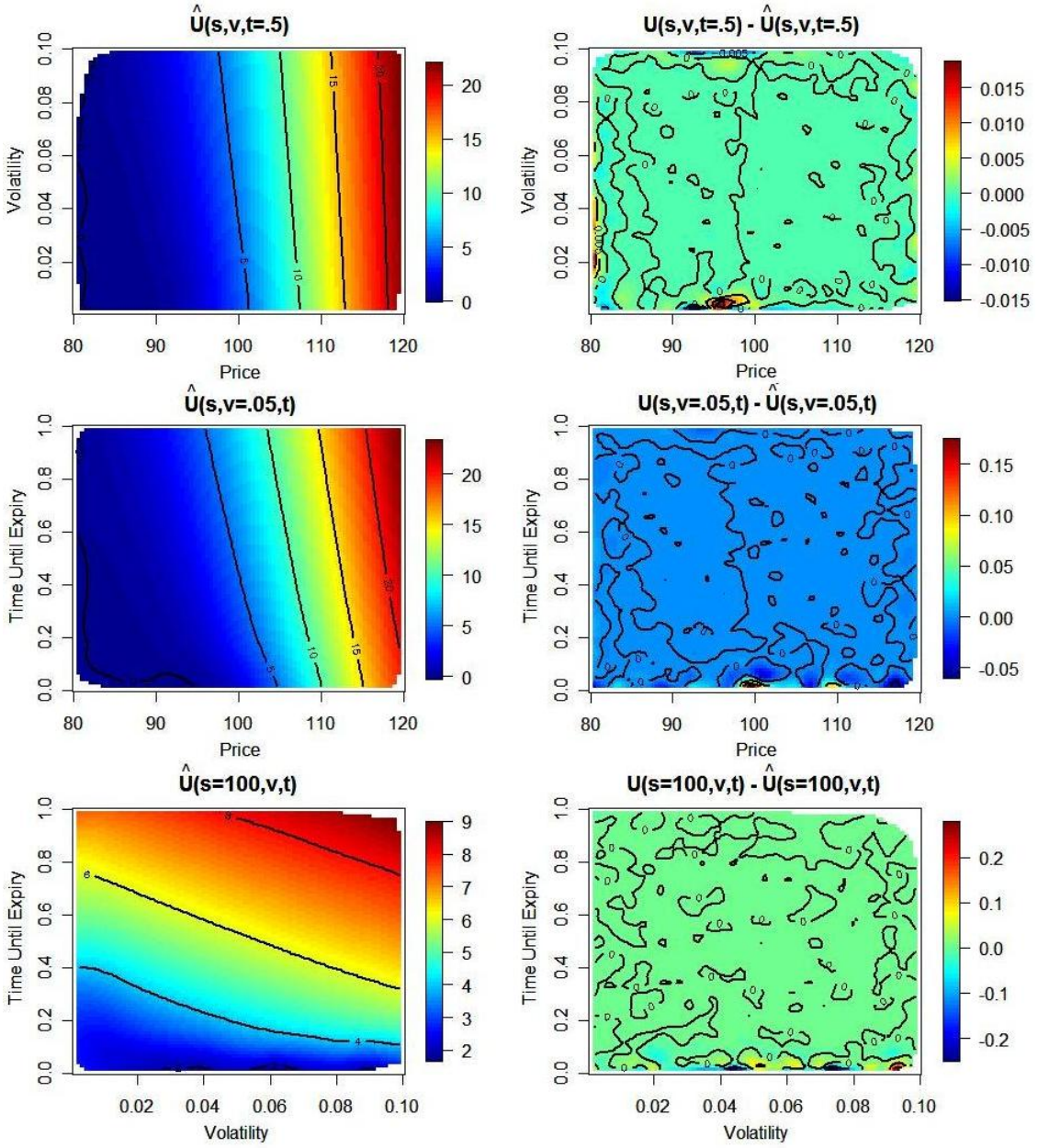

Figure 10: Value surface obtained using thin plate splines under SV.

In this setting, we again have the luxury of a closed-form valuation formula that enables us to examine the performance of the fitted surface as a valuation function. To obtain the surface plots in the right column of Figure 10, we used the three-dimensional fit to predict values for 1000 randomly generated values of two of the three dimensions, the other held fixed.

For those same values, we obtained the closed-form prices and constructed the differences of the closed-form and predicted prices. The surface plots represent these errors as a function of the two relevant dimensions. We see that in all three plots, the fitted surface provides excellent valuation estimates, even near the boundaries where the errors rarely climb above a few cents.

When a valuation mechanism is available, e.g. Monte Carlo simulation in conjunction with LSM, or finite differencing, these results provide a very powerful way of extending a relatively small number of combinations of values of arguments of the unknown valuation function and corresponding valuation estimates to an entire estimated valuation function. Simulating individual values may in fact be time-consuming, yet once the values to which the surface is to be fit are obtained, using that surface as a predictor can lead to reductions in the time required to obtain values for other combinations of the relevant variables used for valuation, in this case, asset price, volatility, and time remaining until expiration. How well the technique works in higher dimensional spaces is left as a topic of future work.

5.2 DECISION RULES FOR DECISION MAKING

In this section we consider the implications of replacing the option payoff functions within LSM with actual option values, which translates to giving the investor the right to trade the option rather than to exercise as soon as it is deemed optimal to do so. This naturally will have no effect on valuation under a risk-neutral measure. We shall investigate, however, any effect it may have on optimal behavior under an assumed real measure and study return distributions when implementing such behavior.

5.2.1 Payoffs From Selling as Opposed to Exercising

We begin by constructing an estimated optimal stopping rule obtained using standard LSM under a risk-neutral measure. At a given time point, the rule is obtained by comparing the known value of immediate exercise with the estimated expected value of continuation.

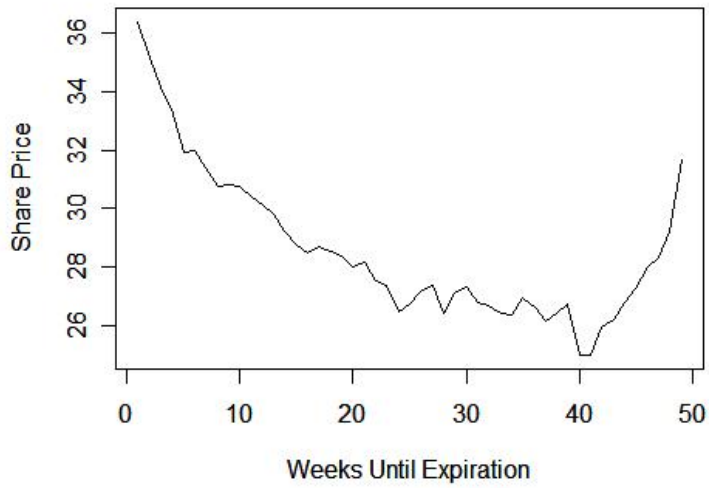

Figure 11: Estimated early exercise boundaries for an American put.

Figure 11 depicts the estimated rule for an American put on a non-dividend paying asset that follows geometric Brownian motion. The annual risk-free interest rate is 0.06, volatility is $\sigma = 0.4$, the strike price of the put is $K = 40$, and the time horizon under consideration is $T = 1$ year. The optimal exercise region is below the estimated boundary displayed in the figure.

When conducting valuation within the LSM framework, at any point in time, the holder of an American stock option faces a decision consisting of two choices; he can do nothing or he can choose to exercise the option. Meanwhile, at a given time t, the option is worth $U(s,t)$, whereas the payoff the investor receives if he chooses to exercise the option is $h(s)$. In reality, $h(s) \leq U(s,t)$, and thus an agent may not choose to exercise the option early, but instead may simply wish to sell that option in the market. There are circumstances under which this is not the case. Certain types of options, for example employee stock options, may not be traded in the market, and instead may only be exercised. Illiquidity may also prevent one from selling options on the market at a fair price. Be this as it may, there are many occasions on which the long option holder will simply choose to sell his options, and we wish to consider any implications stemming from allowing him to do.

Recall within the LSM valuation framework, conditional expected values of continuation were estimated at each allowable early exercise date by regressing discounted pathwise realized future payoffs stemming from exercise against functions of state variables. By comparing this estimated value of continuation, conditional on the current state, with the payoff associated with immediate exercise, an agent is able to decide whether it is optimal to exercise an option early.

For our purposes, however, within an agent's decision space at each time point, we essentially replace *exercise early* with *sell option* and consider how an agent's optimal behavior may change by doing so. By making suitable adjustments to the way LSM is implemented, we can attempt to estimate an agent's optimal behavior under our new assumption that he is allowed to sell options on the market. Those adjustments comprise replacing the dependent variables in the regressions with pathwise discounted realized future cash flows generated by selling the option instead of exercising it. We also must change the values with which these proceeds are compared from the cash flow stemming from immediate exercise to the cash flow stemming from selling the option immediately. Within the usual LSM framework, each of these quantities can be expressed as a known function of s. Meanwhile, for our purposes they stem from an unknown function of both s and t (which is in fact the very function LSM is used to estimate in the first place). Also recall that standard LSM only uses in-the-money sample paths to estimate conditional expectations. This is because the only time an option holder has a choice to make regarding early exercise is when his option is in-the-money. We must remove this constraint since an option holder is free to sell his option at any time he chooses, regardless of whether it is in-the-money, making the decision to sell or continue always relevant.

More formally, at each time point t_l in the set of allowable early exercise dates, LSM assumes that the conditional expected value of continuation

$$G(\omega; t_l) = E_Q[\sum_{j=l+1}^{L} e^{-\int_{t_l}^{t_j} r(\omega,u)du} CF(\omega, t_j; t_l, T)|\mathcal{F}_{t_l}], \tag{5.1}$$

can be written as a countable linear combination of $\mathcal{F}_{t_l}$-measurable basis functions. This condition is satisfied when we are focusing on derivatives with finite-variance random payoffs.

LSM then uses regression to obtain estimates, $\hat{G}_M(\omega; t_l)$, of 5.1 for each t_l using $M < \infty$ basis functions. Finally, an estimated optimal exercise rule (or region) is obtained by comparing, at each time point, $\hat{G}_M(\omega; t_l)$ with the value of immediate exercise, $h(s(t_l))$; an application of the dynamic programming principle implies that it is optimal to exercise the first time at which $\hat{G}_M(\omega; t_l) < h(s(t_l))$.

The first modification we make for our purposes is to replace $CF(\omega, t_j; t_l, T)$ in 5.1 with $CF^*(\omega, t_j; t_l, T)$, the vector of cash flows obtained by following the optimal trading rule at all times s, $t_l < s < T$, conditional on the option not being sold prior to time t_l. Refer to the estimated rule obtained using CF^* as $G^*(\omega; t_l)$. When deciding whether to sell an option, G^* is now compared to $U(s, t_l)$ instead of $h(s(t_l))$, since instead of receiving $h(s(t_l))$ by exercising the option we are receiving $U(s, t_l)$ by selling it. As mentioned above, we also use all simulated sample paths to estimate the conditional expectations since the decision whether to sell the option is always relevant, regardless whether the option is in-the-money. These changes enable us to attempt to estimate the region of state prices, at each time point, within which it is better to sell the option than continue.

One way to adapt standard LSM to our framework would be to simply use LSM to carry out valuation at each necessary point we encounter throughout implementation of the algorithm at which a value is needed. Because LSM is usually carried out on collections of 10,000 paths or more, this is clearly inappropriate and would result in extreme inefficiencies. Our method of incorporating these changes into LSM relies on obtaining estimates of value functions using thin plate splines as discussed in Section 5.1.

The left panel in Figure 12 depicts the results when we tried to estimate optimal trade boundaries for the same American put option described above under the risk neutral measure, Q. We observed boundaries that exhibited a great deal of variation across five implementations of the algorithm. This suggests that an optimal trade rule does not in fact exist under Q, and is what one would expect since the discounted option value is a martingale under Q.

Meanwhile, Girsanov's Theorem implies that the only difference between the risk neutral and real measures when those measures correspond to geometric Brownian motion, is that the value of the drift changes. As we have seen, under the risk neutral measure the drift term of the underlying geometric Brownian motion is the risk free interest rate r. Under the

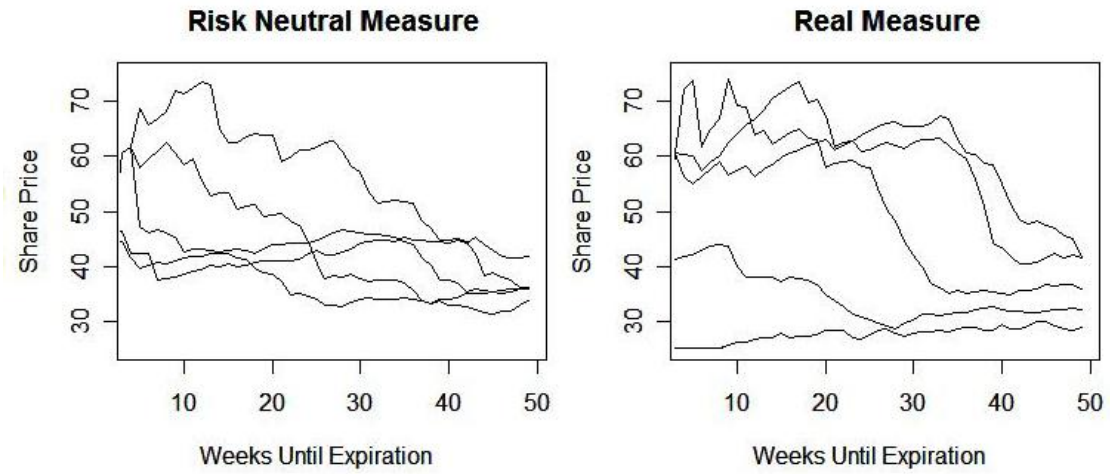

Figure 12: Lack of convergence to any optimal trading rule.

real probability measure, r is replaced by the real drift μ.

Under the real measure, P, the discounted asset value and option value are no longer martingales, therefore we may expect an optimal trade boundary to exist under P. We implemented our modified version of LSM on five collections of simulated asset paths under P, when we assume that $\mu = .12$, for the same American put option as above. The estimated trading rules are depicted in the right panel of Figure 12. Although three of the rules exhibit some similar characteristics, there is still a great deal of variation from rule to rule, suggesting that our methodology fails to find any optimal trading rule, if such a rule exists (since we are considering a long position in an American put and assuming the real rate of return is larger than the risk free rate, one may lobby that the optimal strategy is to abstain from the put position altogether). In fact, we also implemented our modified LSM algorithm using paths simulated under an assumed real rate of return of -.06, and again observed very high variation across all replications.

5.2.2 Returns Using Rules Constructed Under Real Measures

In as much as the estimated optimal stopping rule under the risk neutral measure is required for valuation, it offers little guidance to an agent who already has an established option position because real asset paths evolve according to a real measure that is usually different than the risk neutral measure. Figure 11 depicted an optimal early exercise boundary constructed under the risk neutral measure. Recall in the previous section we mentioned that the only difference between the risk neutral and real measure under geometric Brownian motion is the drift term. We can therefore use standard LSM to construct estimated optimal early exercise rules under the real measure by again simply changing the drift term used to simulate the collection of paths used in the algorithm. Figure 13 depicts rules constructed under the risk neutral measure in which $r = 0.06$ (black), a real measure in which μ is assumed to equal -0.06 (red), a real measure in which μ is assumed to equal 0.0 (blue), and a real measure in which μ is assumed to equal 0.12 (green). As the estimated optimal early exercise region is below the respective boundary, the interpretation is that the smaller we assume the real rate of return to be, the larger the drop in stock price is required to make exercising early optimal.

Although the drift under the real measure is generally unknown, we aim to study the cash flows realized by implementing such rules in practice. We do so by constructing rules under these 4 different assumed real measures and implementing each of those rules on each of 4 different actual real measures. In all cases we assume the risk free rate used for discounting is 0.06. Figure 14 compares the average realized discounted cash flow by following rules constructed under the assumption that the real drift is -0.06 (red), 0.0 (blue), 0.06 (black), and 0.12 (green). For each rule, the average discounted cash flow is calculated when the true underlying drift is -0.06, 0.0, 0.06, and 0.12. We see that the return is the greatest under each measure when we are using the correct rule, and the further our "guess" is from the truth, the worse the rule performs.

The question obviously remains as to which rule to use in the face of uncertainty regarding the underlying real measure. We would like to find a rule that performs reasonably well, regardless of the underlying measure. We chose to construct a rule using standard LSM on

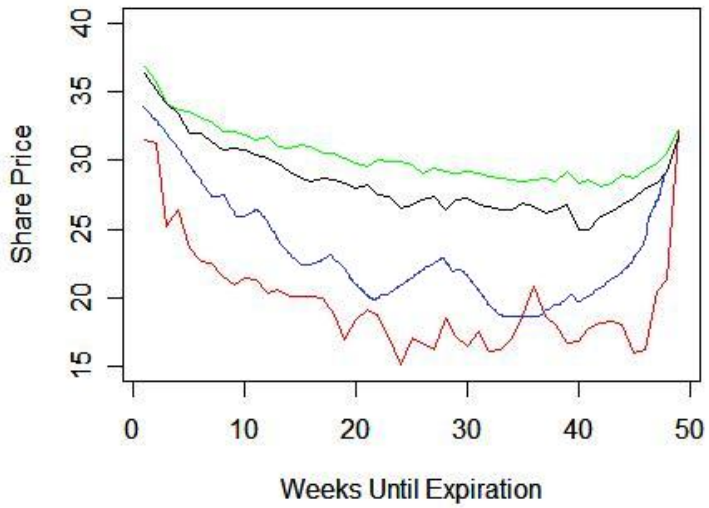

Figure 13: Early exercise rules for an American put.

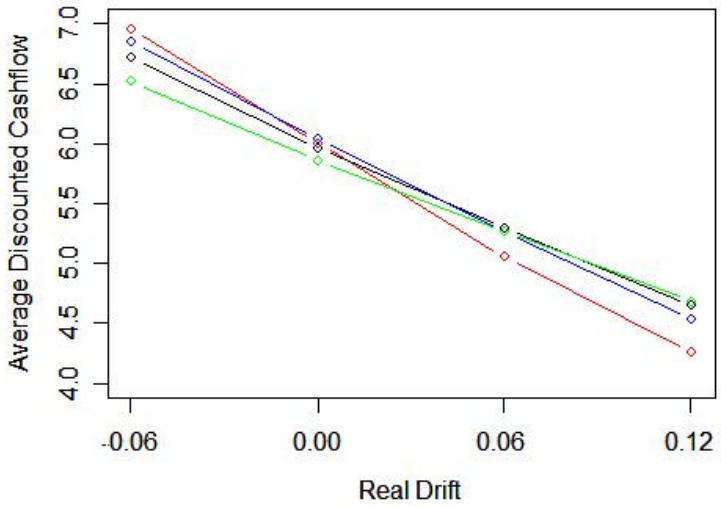

Figure 14: Average cash flows under various rules and real measures.

a collection of simulated paths, but rather than assuming that the real drift of all paths is the same, we randomly chose the drift of the i-th path, μ_i, from the Uniform(-0.06,0.12) distribution. We then applied the rule under the same four measures discussed above.

Table 24: Cashflows from early exercise under various rules.

	$\mu = -0.06$	$\mu = 0.0$	$\mu = 0.06$	$\mu = 0.12$
$\text{RULE}_{-0.06}$	6.958(7.260)	6.002(6.974)	5.050(6.596)	4.250(6.163)
$\text{RULE}_{0.00}$	6.852(6.367)	6.036(6.291)	5.258(6.124)	4.530(5.891)
$\text{RULE}_{0.06}$	6.715(5.727)	5.959(5.724)	5.290(5.676)	4.645(5.534)
$\text{RULE}_{0.12}$	6.521(5.159)	5.917(5.212)	5.266(5.221)	4.680(5.159)
$\text{RULE}_{U(-0.06,0.12)}$	6.802(6.199)	6.028(6.126)	5.285(6.007)	4.573(5.789)

More precisely, we formulated each rule using 100,000 paths simulated under the given dynamic, and then applied the rules to collections of 100,000 paths, each simulated under the given real drift. We repeated this procedure five times and report the average cash flows and corresponding estimated standard deviations of the cash flow distributions in Table 24. The column headings represent the true underlying drift and the row headings represent the drift used to construct each rule. We see that the rule constructed using a varying level of drift performs fairly well. When the real drift is -0.06 or 0.12, our rule ranks third out of five in terms of return, whereas when the real drift is 0.0 or 0.06 our rule ranks second out of five. As the standard deviation of the cash flow distribution decreases with the drift used to construct the rule, it is not surprising that the standard deviation corresponding to our rule ranks consistently in the middle across all real drifts.

5.3 FUTURE WORK

As a focus of other future work, we shall examine various option trading strategies such as **covered call** writing. The writer of a covered call has a long position in one share of the underlying asset and sells one short call option against that share. This very practical strategy is commonly used as a hedge against the risk associated with holding the long position in

the stock, and also to generate the cash flow associated with selling the call. Furthermore, as time progresses and the prices of the underlying asset and the option fluctuate, the agent at some point may wish to close the short position in the option, and concurrently sell another option, having a different strike price and/or expiration date, against his long share of the asset. For instance, if the option initially shorted is at-the-money, and the underlying asset value rises or falls significantly, much of the time value component of that option's value will become depleted. If it is the strategy of the agent to capitalize on time value deterioration of the short option, he may wish to close the existing option and short another closer to the money. As the cash flows generated using this strategy depend on both the time and state in which options are sold, the question the agent needs to consider is twofold: 1) when is it optimal to close the position in one option, and 2) what other option is it optimal to concurrently sell? To address these questions, one possible approach is to consider how the partial derivatives of an option pricing formula change with respect to time and fluctuations in the value of the underlying asset, and also with the underlying real measure.

Define an **option series** to be all options of the same type on the same underlying asset that expire on the same date. Introduce the notation $\hat{U}^T(s,t) = (U^T_{C_1}(s,t), \ldots, U^T_{C_R}(s,t))$ to represent the prices at time t of a series of options when the value of the underlying asset is $s(t)$, all of which expire at time T and have strike prices $C_1, \ldots, C_R$, where R is the number of different strike prices in the series. Also introduce the notation $T_1, \ldots, T_H$, where $T_1 < T_{2<} \cdots < T_H = T$. These T's represent the expiration dates of all series of options under consideration. As before, we can approximate the continuous interval of dates on which trading activity may occur between successive expiration dates with the discrete set of time points $t_{h,0}, t_{h,1}, \ldots,$ and $t_{h,L}$, where $t_{h,0} < t_{h,1} < \ldots < t_{h,L} = T_h$. For simplicity, we shall assume that the $t_{h,l}$'s are evenly spaced. The case $H = 2$ and $L = 3$ is illustrated in Figure 15.

Let $\hat{\phi}^{T^*}(t)$ represent an adapted R-dimensional vector process consisting of component processes $\phi^{T^*}_{C_r}(t)$, which represent the number of options in series T^* with strike price C_r that we own at time t. We can express the change in value of a portfolio from time 0 to time T

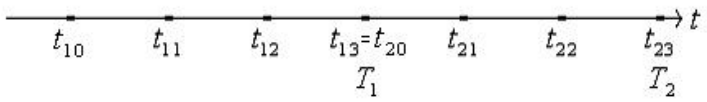

Figure 15: Discrete set of exercise dates.

associated with the option shorting strategy $\phi(t) = (\hat{\phi}^{T_1}(t), \ldots, \hat{\phi}^{T_H}(t))$ as

$$V_\phi(T) - V_\phi(0) =$$

$$\sum_{h=1}^{H}[\sum_{h'=h}^{H}\sum_{l=0}^{L}\sum_{r=1}^{R}\phi_{C_r}^{T_{h'}}(t_{h,l})[U_{C_r}^{T_{h'}}(s, t_{h',l+1}) - U_{C_r}^{T_{h'}}(s, t_{h',l})]].$$

Let Φ represent the collection of all option trading strategies ϕ. Then it should be the goal of a rational agent to maximize $e^{-r(T-t)}E[V_\phi(T)|\mathcal{F}]$ over Φ for each t. Under the risk neutral measure this quantity is a martingale, however we are able to seek strategies under real measures with the goal of maximizing the expectation. Similarly, an agent may wish to minimize $\mathrm{Var}[V_\phi(T)|\mathcal{F}]$.

When we hold an asset long against the options shorted using strategy ϕ, the change in value of the portfolio during the interval $(0, T)$ also consists of any change in value associated with these assets. We shall assume that whatever asset A is held long against options shorted at the beginning of the strategy is in fact held throughout the duration of the strategy. For our considerations, A can be either a share of stock as in covered call writing, or another option on the same underlying asset as in horizontal and diagonal spread writing. Then the change in equity of the portfolio during the interval $(0, T)$ is

$$[V_\phi(T) - V_\phi(0)] + [A(T) - A(0)]$$

We shall begin by placing suitable restrictions on the types of trading that can occur, which are often the types of restrictions associated with margin trading requirements imposed by the Securities Exchange Commission. This essentially translates to restricting the state

space of ϕ. Impose the restriction that exactly one option must be shorted against the long asset A at each $t \in (0,T)$. If a short option expires worthless on T_h, then another option is written at that time having some later expiration date. If an option is held until expiration, at which time it is in-the-money, the payoff is awarded to the holder of the option and another option is written at that time having some later expiration date. No restriction is placed on either the strike price or the expiration date of any shorted option, other than that it expires no later than T. In summary, whenever the next expiration date is T_h, this restricted state space of ϕ is $\{0,1\}^{H\times R}$, subject to the constraint

$$\sum_{h'=h}^{H}\sum_{r=1}^{R}\phi_{C_r}^{T_{h'}}(t)=1.$$

In covered call writing, we shall assume that at time 0 we are long exactly one share of stock, whose price process is given by $A(t)=S(t)$. We wish to write call options as described above against this share of stock during $(0,T)$. A guidepost we can use in formulating our strategies pertains to the partial derivatives of the options under consideration. One advantage of selling a call option is that the time value component of the option erodes at a rate much higher than the risk free interest rate whenever the strike price of the call is close to the asset price. The partial derivative of an option valuation formula with respect to time is steepest for options which are at or near-the-money. Furthermore, due to the convexity of a pricing formula, with state held fixed the time value component of an option will erode at an increasingly higher rate as expiration approaches. These reasons suggest that an effective strategy whenever the next expiration date is $T_{h'}$ may be to choose

$$\phi_{C_r}^{T_{h'}}(t_{h,l})=I_{[\omega',h=h',r\neq r']}(\omega,h,r) \tag{5.2}$$

where $h\in\{h',\ldots,H\}$, $r\in\{1,\ldots,R\}$, and

$$\omega'=[\omega:\frac{U_{C_r}^{T_{h'}}(s,t_{h,l+1})}{U_{C_r}^{T_{h'}}(s,t_{h,l})}>\frac{U_{C_{r'}}^{T_{h'}}(s,t_{h,l+1})}{U_{C_{r'}}^{T_{h'}}(s,t_{h,l})}].$$

In the case of geometric Brownian motion with r and σ both assumed known and constant, we have the luxury of being able to calculate these ratios of partial derivatives. In this case, we shall investigate how this suggested strategy performs for various levels of volatility by

comparing simulated returns with the risk free rate. Under geometric Brownian motion, we can also investigate other strategies which depend on an option's delta in addition to theta.

We shall also consider corresponding strategies for other price dynamics where we lack convenient access to an explicit option pricing formula and the required partial derivatives. In this case, all of these quantities can be simulated, after which we can formulate the option shorting strategy. The simple method we propose is to simulate a price surface for an option using LSM and the thin plate spline approach of Section 5.1. From this surface we are able to obtain estimates of the required partial derivatives of the pricing formula. For example, an estimate of an option's theta at time $t_{h,k}$ when the value of the underlying security is s is

$$\frac{\hat{U}_{C_r}^{T_h}(s, t_{h,l+1}) - \hat{U}_{C_r}^{T_h}(s, t_{h,l})}{t_{h,l+1} - t_{h,l}}.$$

After obtaining estimates of the required partial derivatives, we can formulate the option trading strategy using Equation 5.2.

5.4 CONCLUSIONS

Much of this dissertation has focused on improving the efficiency of Monte Carlo estimators of stock option prices. We developed estimators for both European and American option values that outperformed existing methods. We also provided techniques to increase the efficiency of the Least Square Monte Carlo algorithm. In the final chapter, we presented the effectiveness of using thin plate spline methodology in conjunction with existing valuation techniques to produce an option valuation surface. Our results are useful to the financial engineering industry, in which accurate and fast valuation techniques are of great importance.

APPENDIX A

GLOSSARY OF FINANCIAL TERMINOLOGY

American option - an option contract that can be exercised at any time between the date of purchase and the expiration date

arbitrage - the simultaneous buying and selling of a security at two different prices in two different markets, resulting in profits without risk; perfectly efficient markets present no arbitrage opportunities (perfectly efficient markets seldom exist, but, arbitrage opportunities are often precluded because of transactions costs)

asset price dynamics - a theoretical model governing the evolution of an asset's value through time; common dynamics are geometric brownian motion, geometric brownian motion with poisson jumps, jump diffusion processes, etc.

call option - an option that gives the holder the right to buy the underlying security

contingent claim - a claim that can be made only if a specified outcome occurs

delta - measures an option's first order (linear) sensitivity to a change in the value of the underlying asset

derivative security - a financial security such as an option whose value is derived in part from the value and characteristics of another security, the underlying asset

dividend - a portion of a company's profit paid to common and preferred shareholders

efficient market hypothesis - states that all relevant information is fully and immediately reflected in a security's market price, thereby assuming that an investor will obtain an equilibrium rate of return; in other words, an investor should not expect to earn an abnormal return (above the market return) through either technical analysis or fundamental analysis

ex-dividend - literally means "without dividend"; the buyer of shares when they are quoted ex-dividend is no longer entitled to receive a declared dividend

exercise - implementing the right by the holder of an option to buy (in the case of a call) or sell (in the case of a put) the underlying security

European option - an option that may be exercised only on the expiration date

exotic option - refers to options that are more complex than simple puts or call options; for example, a caput is a call option on a put option

expiration - the time after which the terms of an option contract are no longer valid

in-the-money - a put option that has a strike price higher than the underlying asset price, or a call option with a strike price lower than the underlying asset price

intrinsic value - the amount by which an option is in-the-money

option - gives the buyer the right, but not the obligation, to buy or sell an asset at a set price over a prescribed period of time

out-of-the-money - a put option that has a strike price lower than the underlying asset price, or a call option with a strike price higher than the underlying asset price

put option - an option granting the right to sell the underlying security

risk-free rate - the rate earned on a riskless asset, e.g. U.S. government bonds

strike price - the price at which the security underlying an options contract may be bought or sold

time value - the difference between the price of an option and the option's intrinsic value

underlying asset - the security that an option gives the option holder the right to buy or sell

volatility - a measure of risk based on the standard deviation of the asset return; volatility is a variable that appears in option pricing formulas, where it denotes the volatility of the underlying asset return from now until the expiration of the option

APPENDIX B

LIST OF COMMON NOTATION

$X(t)$	a stochastic process at time t; $S(t)$, $V(t)$, and $W(t)$ represent a price process, volatility process, and Brownian motion, respectively
$x(t)$	the realized value of a process at time t; *mutatis mutandis* $s(t)$, $v(t)$, and $w(t)$
$u(s,t)$	the value of a European option with time $T-t$ remaining until expiry, when the asset is worth s at time t
$U(s,t)$	the value of an American option with time $T-t$ remaining until expiry, when the asset is worth s at time t
r	the rate of return on a risk-free asset
K	the strike price of an option
$h(\cdot)$	payoff function associated with an option
T	the time before which all economic activity under consideration transpires; usually associated with the expiration date of an option
ϕ	a trading strategy
L	the number of time points at which early exercise is allowed within the LSM algorithm
M	the number of basis functions used when implementing the LSM algorithm

APPENDIX C

R CODE FOR PRIMARY ALGORITHMS

Exact simulation under the SV model:

```
>mainExactSim
function (s,strike,u,t,vu,r,kap,theta,sigv,rho,upLim,h,lastBes,n){
##s = beginning asset price
##strike = strike price of option
##u = initial time
##t = end time
##vu = initial variance
##r = risk free rate
##kap = risk free force of mean reversion of variance process
##theta = long run mean of volatility process
##sigv = volatility of variance process
##rho = correlation between asset path and variance path
##upLim = the largest value at which we evaluate the cf
##h = the grid width in the trapezoidal approximation to
##the fourier inversion of the distribution
##lastBes = number of terms used to calculate the bessel function
##in method1
##vt given vu
```

```
d<-(4*theta*kap)/sigv^2
ncp<-(4*kap*exp(-kap*(t-u))*vu)/(sigv^2*(1-exp(-kap*(t-u))))
vt<-((sigv^2*(1-exp(-kap*(t-u))))/(4*kap))*rchisq(n,d,ncp)
##integVDs given vu and vt
integVDs<-numeric(n)
grid<-seq(0,upLim,by=h)
gam<-sqrt(kap^2-(2*sigv^2*1i*grid))
first<-(gam*exp(-.5*(gam-kap)*(t-u))*(1-exp(-kap*(t-u))))/(kap*(1-exp(-gam*(t-u))))
for(q in 1:n){
second<-exp(((vu+vt[q])/sigv^2)*((kap*(1+exp(-kap*(t-u))))/
(1-exp(-kap*(t-u)))-(gam*(1+exp(-gam*(t-u))))/(1-exp(-gam*(t-u)))))
besDenom<-rep(besselI(((sqrt(vu*vt[q])*(4*kap*exp(-.5*kap*(t-u))))/
(sigv^2*(1-exp(-kap*(t-u))))),(.5*d-1)),length(grid))
besNum<-method1ExactSim(rep((.5*d-1),length(grid)),((sqrt(vu*vt[q])*
(4*gam*exp(-.5*gam*(t-u))))/(sigv^2*(1-exp(-gam*(t-u))))),lastBes)
charFunc<-first*second*(besNum/besDenom)
reCharFunc<-Re(charFunc)
sam<-runif(1)
toFindRoot<-function (x)(((h*x)/pi)+(2*h/pi)*
sum((sin(grid*x)*reCharFunc/grid)[2:length(grid)])-sam)
integVDs[q]<-uniroot(toFindRoot,low=0,up=50*(vt[q]+vu)*(t-u),tol=.0001)[1]
}
integVDs<-as.numeric(integVDs)
##integRootVDws1
integRootVDws1<-(1/sigv)*(vt-vu-kap*theta*(t-u)+kap*integVDs)
##s
muVect<-log(s)+(r*(t-u)-.5*integVDs+rho*integRootVDws1)
varVect<-(1-rho^2)*integVDs
sSamp<-exp(muVect+sqrt(varVect)*rnorm(n,0,1))
return(mean(pmax(sSamp-strike,0)))
```

```
}

>method1ExactSim
function (v,z,last){
##v = order of the bessel function
##z = (complex) argument of the bessel function
##last = number of components in the power series
##expansion used to calculate the bessel function
j<-0:last
out<-complex(length(v))
for(i in 1:length(v)){
contrib<-((.25*z[i]^2)^j)/(factorial(j)*gamma(v[i]+j+1))
out[i]<-((.5*z[i])^v[i])*sum(contrib)
}
##correct for discontinuities
jump<-0
for (i in 3:length(out)){
{if((abs(Arg(out[i])-Arg(out[i-1]))>5*abs(Arg(out[i-1])
-Arg(out[i-2])))|(abs(Arg(out[i])-Arg(out[i-1]))
>5*abs(Arg(out[i-1])-Arg(out[i-2])))){
jump<-rbind(jump,i)
}
if(length(jump)>1)
for(m in 2:length(jump)){
out[jump[m]:length(out)]<-out[jump[m]:length(out)]
*exp(2*v[1]*pi*1i)
}
return(out)
}
```

Least Squares Monte Carlo Algorithm for a Put Option Under GBM:

```
##main LSM
function(s,strike,sig,riskFreeRate,realRate,n,m){
##s is the stock price at the beginning of the period
##strike is the strike price of the option under consideration
##sig is the instantaneous standard deviation of the GBM
##riskFreeRate
##realRate
##n is the number of paths used in the simulation
##m is the number of points we allow for early exercise (usually 50/year)

sMat<-method2(s,sig,realRate,n,m)
cashFlowMat<-as.matrix(ifelse(sMat[,m]<strike,strike-sMat[,m],0))
m<-m-1
cashFlowMat<-method1(strike,riskFreeRate,m,sMat,cashFlowMat)
counter<-apply(cashFlowMat,1,function(x) ifelse(max(x)==0,1,which(x>0)[1]))
  discCashFlow<-cashFlowMat[cbind(1:n,counter)]*exp(-counter*riskFreeRate/50)
return(mean(discCashFlow))
}

##method1 -- LSM recursion
function(strike,riskFreeRate,m,sMat,cashFlowMat){
if(m==0){return(cashFlowMat)}
else{
numInMoney<-sum(sMat[,m]<strike)
sy<-numeric(2*numInMoney)
dim(sy)<-c(numInMoney,2)
inMoney<-which(sMat[,m]<strike)
reducedCashFlowMat<-cashFlowMat[inMoney,]
reducedCashFlowMat<-as.matrix(reducedCashFlowMat)
```

```
sy[,1]<-sMat[inMoney,m]
counter<-apply(reducedCashFlowMat,1,function(x)
ifelse(max(x)==0,1,which(x>0)[1]))
sy[,2]<-reducedCashFlowMat[cbind(1:numInMoney,counter)]
*exp(-counter*riskFreeRate/50)
s<-sy[,1]
sSq<-sy[,1]^2
sCu<-sy[,1]^3
sFo<-sy[,1]^4
sFi<-sy[,1]^5
sSi<-sy[,1]^6
sSe<-sy[,1]^7
lmFit<-lm(sy[,2]~s+sSq+sCu+sFo+sFi+sSi+sSe)
dfForLm<-data.frame(s=sMat[,m],sSq=sMat[,m]^2,sCu=sMat[,m]^3,
sFo=sMat[,m]^4,sFi=sMat[,m]^5,sSi=sMat[,m]^6,sSe=sMat[,m]^7)
theFit<-predict(lmFit,dfForLm)
cashFlowNew<-ifelse((sMat[,m]<strike)&((strike-sMat[,m])>theFit),
strike-sMat[,m],0)
cashFlowMat<-cbind(cashFlowNew,cashFlowMat)
return(method1(strike,riskFreeRate,m-1,sMat,cashFlowMat))
}
}

##method2 -- Generate GBM paths
function (s,sig,realRate,n,m){
sMat<-numeric(n*(m+1))
dim(sMat)<-c(n,(m+1))
sMat[,1]<-s
for (j in 1:m){
a<-rnorm(n)
```

```
sMat[,j+1]<-sMat[,j]*exp(((mu-(sig^2/2))/50)+(sig*sqrt(1/50)*a))
}
return(sMat[,2:(m+1)])
}
```

Generate Faure numbers:

```
>mainFaureGen
function(n,p,s){
##n = number of Faure numbers to return
##p = prime number to use in expansion
##s = dimension of Faure numbers
dimNCR<-trunc(log(n,base=p))+1
nCRMat<-numeric((dimNCR)^2)
dim(nCRMat)<-c((dimNCR),(dimNCR))
for (i in 1:dimNCR){
for (j in 1:dimNCR){
nCRMat[i,j]<-choose(j-1,i-1)
}
}
baseRep<-numeric(dimNCR*n)
dim(baseRep)<-c(n,dimNCR)
intList<-seq(1:n)
for (j in dimNCR:1){
baseRep[,j]<-trunc((intList/p^(j-1))%%p)
}
faureMat<-numeric(n*s)
dim(faureMat)<-c(n,s)
for (i in 1:dimNCR){
faureMat[,1]<-faureMat[,1]+baseRep[,i]*p^-i
```

```
}
for (i in 1:n){
fromBaseRep<-baseRep[i,]
for (j in 2:s){
nextBaseRep<-(nCRMat%*%fromBaseRep)%%p
for (k in 1:dimNCR){
faureMat[i,j]<-faureMat[i,j]
+nextBaseRep[k]*p^-k
}
fromBaseRep<-nextBaseRep
}
}
return(faureMat)
}
```

Heston Option Valuation Formula

```
function(s,k,v,r,rfkappa,rftheta,sigv,rho,tau){
##s = price of underlying asset
##k = strike price of option
##v = volatility of underlying asset
##r = risk free interest rate
##rfkappa = risk neutral force of reversion
##rftheta = risk neutral mean variance
##sigv = volatility of variance process
##rho = instantaneous correlation between price and volatility
##tau = time until expiration
u1<-.5
u2<--.5
a<-rfkappa*rftheta
```

```
b1<-rfkappa-(rho*sigv)
b2<-rfkappa
fee<-seq(.00001,100.00001,length=5001)
d1<-sqrt(((rho*sigv*fee*1i-b1)^2)-(sigv^2*(2*u1*fee*1i-fee^2)))
d2<-sqrt(((rho*sigv*fee*1i-b2)^2)-(sigv^2*(2*u2*fee*1i-fee^2)))
g1<-(b1-rho*sigv*fee*1i+d1)/(b1-rho*sigv*fee*1i-d1)
g2<-(b2-rho*sigv*fee*1i+d2)/(b2-rho*sigv*fee*1i-d2)
C1<-(r*fee*1i*tau)+((a/sigv^2)*(tau*(b1-rho*sigv*fee*1i+d1)
-2*log((1-g1*exp(d1*tau))/(1-g1))))
C2<-(r*fee*1i*tau)+((a/sigv^2)*(tau*(b2-rho*sigv*fee*1i+d2)
-2*log((1-g2*exp(d2*tau))/(1-g2))))
D1<-((b1-rho*sigv*fee*1i+d1)*(1-exp(d1*tau))) / (sigv^2*(1-g1*exp(d1*tau)))
D2<-((b2-rho*sigv*fee*1i+d2)*(1-exp(d2*tau))) / (sigv^2*(1-g2*exp(d2*tau)))
f1<-exp(C1+(D1*v)+(1i*fee*log(s)))
f2<-exp(C2+(D2*v)+(1i*fee*log(s)))

##correct for any discontinuities in f1 and f2
jump1<-0
jump2<-0
for (i in 3:(length(fee)-1))
{
if(sqrt((Re(f1[i])-Re(f1[i-1]))^2+(Im(f1[i])-Im(f1[i-1]))^2) >
5*sqrt((Re(f1[i-1])-Re(f1[i-2]))^2+(Im(f1[i-1])-Im(f1[i-2]))^2))
{jump1<-rbind(jump1,i)}

if(sqrt((Re(f2[i])-Re(f2[i-1]))^2+(Im(f2[i])-Im(f2[i-1]))^2) >
5*sqrt((Re(f2[i-1])-Re(f2[i-2]))^2+(Im(f2[i-1])-Im(f2[i-2]))^2))
{jump2<-rbind(jump2,i)}
}
if(length(jump1)>1)
```

```
for(m in 2:length(jump1))
{f1[jump1[m]:length(f1)]<-f1[jump1[m]:length(f1)]*exp(1i*(-(Arg(f1[jump1[m]])
-Arg(f1[jump1[m]-1]))+(Arg(f1[jump1[m]+1])-Arg(f1[jump1[m]])))))}
if(length(jump2)>1)
for(m in 2:length(jump2))
{f2[jump2[m]:length(f2)]<-f2[jump2[m]:length(f2)]*exp(1i*(-(Arg(f2[jump2[m]])
-Arg(f2[jump2[m]-1]))+(Arg(f2[jump2[m]+1])-Arg(f2[jump2[m]])))))}

##trapezoidal rule approximation
integrand1<-Re((exp(-1i*fee*log(k))*f1)/(1i*fee))
integrand2<-Re((exp(-1i*fee*log(k))*f2)/(1i*fee))
P1<-.5+((1/100)*(sum(integrand1)+sum(integrand1[2:5000])))/pi
P2<-.5+((1/100)*(sum(integrand2)+sum(integrand2[2:5000])))/pi
return((s*P1)-(k*exp(-r*tau)*P2))
}
```

BIBLIOGRAPHY

[1] Abramowitz, M. and Stegun, I.A. (1972) *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. National Bureau of Standards, Washington D.C.

[2] Akgiray, V., and Booth, G. (1986) Stock Price Processes With Discontinuous Time Paths: An Empirical Examination. *The Financial Review* **21**, 162-184.

[3] Bates, David S. (1996) Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options. *The Review of Financial Studies* **9**, 69-107.

[4] Billingsley, P. (1995) *Probability and Measure, 3rd edition*. Wiley, New York.

[5] Black, F. and Scholes, M. (1973) The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* **81**, 635-654.

[6] Boyle, Phelim P. (1977) Options: A Monte Carlo Approach. *Journal of Financial Economics* **4**, 323-338.

[7] Broadie, M. and P. Glasserman (1997) Pricing American-Style Securities Using Simulation. *Journal of Economic Dynamics and Control* **21**, 1323-1352.

[8] Broadie, M. and Kaya, O. (2006) Exact Simulation of Stochastic Volatility and Other Affine Jump Diffusion Processes. *Operations Research* **54**, 217-231.

[9] Clement, E., Lamberton, D., and Protter, P. (2002) An Analysis of the Longstaff-Schwartz Algorithm for American Option Pricing. *Finance and Stochastics* **6**, 449-471.

[10] Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985) A Theory of the Term structure of Interest Rates. *Econometrica* **53**, 385-408.

[11] Cox, John C. and Ross, Stephen A. (1976) The Valuation of Options For Alternative Stochastic Processes. *Journal of Financial Economics* **3**, 145-166.

[12] Cox, John C., Ross, Stephen A., and Rubenstein, Mark (1979) Option Pricing: A Simplified Approach. *Journal of Financial Economics* **7**, 229-263.

[13] Duffie, D., Singleton, K., and Pan, J. (2000) Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica* **68**, 1343-1376.

[14] Egloff, D. (2005) Monte Carlo Algorithms for Optimal Stopping and Statistical Learning. *The Annals of Applied Probability* **15**, 1396-1432.

[15] Glasserman, P. (2004) *Monte Carlo Methods in Financial Engineering*. Springer.

[16] Green, P.J., and Silverman, B.W. (1994) *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. Chapman and Hall, London.

[17] Harrison, J. M. and Kreps, D. M. (1979) Martingales and Arbitrage in Multiperiod Securities Markets. *Journal of Economic Theory* **20**, 381-408.

[18] Harrison, J. M. and Pliska, S. R. (1981) Martingales and Stochastic Integrals in the Theory of Continuous Trading. *Stochastic Processes and Their Applications* **11**, 215-260.

[19] Harrison, J. M. and Pliska, S. R. (1983) A Stochastic Calculus Model of Continuous Trading: Complete Markets. *Stochastic Processes and Their Applications* **15**, 313-316.

[20] Hastie, T., Tibshirani, R., and Friedman, J. (2001) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York.

[21] Heston, S. (1993) A Closed-Form Solution of Options With Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies* **6**, 327-343.

[22] Hull, John and White, Alan. (1990) Valuing Derivative Securities Using the Explicit Finite Difference Method. *Journal of Financial and Quantitative Analysis* **25**, 87-100.

[23] Johnson, Herb and Shanno, David. (1987) Option Pricing When the Variance is Changing. *Journal of Financial and Quantitative Analysis* **22**, 143-151.

[24] Joy, Corwin, Boyle, Phelim, and Tan, Ken Seng. (1996) Quasi-Monte Carlo Methods in Numerical Finance. *Management Science* **42**, 926-938.

[25] Karatzas, I. and Shreve, S.E. (1988) *Brownian Motion and Stochastic Calculus*. Springer, Berlin.

[26] Karlin, S. and Taylor, H.M. (1975) *A First Course in Stochastic Processes*. Academic Press, New York.

[27] Karlin, S. and Taylor, H.M. (1981) *A Second Course in Stochastic Processes*. Academic Press, New York.

[28] Lintner, J. (1965) The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *Review of Economics and Statistics* **47**, 13-37.

[29] Longstaff, Francis A. and Schwartz, Eduardo S. (2001) Valuing American Options by Simulation: A Simple Least-Squares Approach. *The Review of Financial Studies* **14**, 113-147.

[30] Merton, R. C. (1973) Theory of Rational Option Pricing. *Bell Journal of Economics and Management Science* **4**, 141-183.

[31] Merton, R. C. (1976) Option Pricing when Underlying Stock Returns are Discontinuous. *Journal of Financial Economics* **3**, 125-144.

[32] Merton, R. C. and P. A. Samuelson (1974) Fallacy of the Log-Normal Approximation of Optimal Portfolio Decision-Making Over Many Periods. *Journal of Financial Economics* **1**, 67-94.

[33] Moreno, M. and Navas, J. (2003) On the Robustness of Least-Squares Monte Carlo (LSM) for Pricing American Derivatives. *Review of Derivatives Research* **6**, 107-128.

[34] Neiderreiter, H. (1992) *Random Number Generation and Quasi-Monte Carlo Methods*. SIAM, Philadelphia, Pa.

[35] Pitman, J. and Yor, M. (1982) A Decomposition of Bessel Bridges. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **59**, 425-457.

[36] Raymar, S. and Zwecher, M. (1997) A Monte Carlo Valuation of American Call Options on the Maximum of Several Stocks. *Journal of Derivatives* **5**, 7-23.

[37] Schwartz, E. S. (1977) The Valuation of Warrants: Implementing a New Approach. *Journal of Financial Economics* **4**, 79-93.

[38] Sharpe, W.F. (1964) Capital Asset Prices: A Theory of Equilibrium Under Conditions of Risk. *Journal of Finance* **19**, 425-442.

[39] Shreve, S. (2004) *Stochastic Calculus for Finance I: The Binomial Asset Pricing Model*. Springer, New York.

[40] Shreve, S. (2004) *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer, New York.

[41] Tavella, D. (2002), *Quantitative Methods in Derivative Pricing*. John Wiley & Sons.

[42] Wahba, G. (1990), *Spline Models for Observational Data*. SIAM, Philadelphia.