Alternative Price Processes for Black-Scholes: Empirical Evidence and Theory *

Samuel W. Malone

April 19, 2002

*This work is supported by NSF VIGRE grant number DMS-9983320.

1 Introduction

This paper examines the price process assumption of the Black-Scholes equation for pricing options on an empirical and analytical level. The material is organized into six primary sections. The first section presents several key results in probability theory, including (most importantly) Itô's Lemma, which is used in the derivation of Black-Scholes. In the second section, we derive the Black-Scholes equation and discuss its assumptions and uses. For the sake of simplicity, we restrict our attention to the case of a European call option, but the analyses herein can be extended to other types of derivatives as well. In the section following this, we examine the assumption made in the Black-Scholes methodology that security prices follow a geometric Brownian motion and discuss empirical evidence to the contrary. Next, we present several analytically useful alternatives for the price process, including alternative diffusions, jump processes, and a few models suggested by the empirical literature. In this section, several processes are described precisely but never explicitly used to obtain option pricing formulas; these problems will hopefully be the focus of future research. The final major section discusses the general theory of option pricing for alternative stochastic processes and applies this theory to some of the candidate processes that have been proposed in the literature. The last section concludes.

2 Probability Background and Ito's Lemma

This section is meant as a refresher and an overview of concepts in probability theory that are related to several of the topics discussed in this paper. As such, most of the results are presented very suffinctly, and the initiated reader can skip directly to section 3. Unless otherwise stated, the material discussed in this section is drawn from Probability by Breiman (1968).

2.1 Stochastic Processes, the Markov Property, and Diffusions

Following Lawler (1995), a stochastic process is a random process evolving with time. In particular, a stochastic process is a collection of random variables $X(t)$ indexed by time. In this paper, time is always a subset of $[0, \infty)$, the nonnegative real numbers. Thus, all stochastic processes considered herein are "continuous time" stochastic processes. The random variable $X(t)$ takes values in a set called the state space, which in this paper is usually the continuous space of prices or returns. The space of prices is the nonnegative real numbers, and the space of returns is the real line. A large class of stochastic processes have the property that the change at time t is determined by the value of the process at time t and not by the values at times before t. Such processes are known as Markov processes (Lawler, 1995). When the transition probabilities for a Markov process do not depend on time, we refer to the process as a time-homogeneous Markov process. All processes considered in this paper are of this form. Throughout, we use the term "diffusion" to denote a continous time, continuous state space, Markov process whose sample paths are continuous. Not all the processes discussed are diffusions, however; we also use Poisson processes, which are continuous time, continuous state space Markov processes with discontinuous sample paths.

2.2 Convergence in Distribution

Definition 2.1. We say that X_n converges to X in distribution, denoted $X_n \xrightarrow{\mathcal{D}} X$, if $F_n(x) \to F(x)$ at every point $x \in C(F)$, the set of continuity points of F. We also write in this case $F_n \xrightarrow{\mathcal{D}} F$.

2.3 Characteristic Functions

Definition 2.2. Given a distribution function $F(x)$, its characteristic function $\phi(u)$ is a complex-valued function defined on $\Re^1$ by

$$\phi(u) = \int e^{iux} F(dx).$$

If F is the distribution function of the random variable X, then equivalently,

$$\phi(u) = E(e^{iuX}).$$

Any characteristic function $\phi(u)$ has the following properties:

(i) $\phi(0) = 1$,
(ii) $|\phi(u)| \leq 1$,
(iii) $\phi(u)$ is uniformly continuous on $\Re^1$,
(iv) $\phi(-u) = \bar{\phi}(u)$.

The following important theorem deals with convergence in distribution.

Theorem 2.3. *(The Continuity Theorem). If F_n are distribution functions with characteristic functions $\phi_n(u)$ such that*

(a) $\lim_{n\to\infty} \phi_n(u)$ *exists for every* u
(b) $\lim_{n\to\infty} \phi_n(u) = \phi(u)$ *is continuous at* $u = 0$,

then there exists a distribution function F such that $F_n \xrightarrow{\mathcal{D}} F$ and $\phi(u)$ is the characteristic function of F.

The next statement is an immediate consequence of this result:

Corollary 2.4. *Let F_n be distribution functions with characteristic functions ϕ_n. If there is a distribution function F with characteristic function ϕ such that*

$$lim_{n\to\infty}\phi_n(u) = \phi(u)$$

for every u, then

$$F_n \xrightarrow{\mathcal{D}} F.$$

Every characteristic function corresponds to a unique distribution function. Sometimes it is useful to know how, given a characteristic function, to find the corresponding distribution function. While most important facts about characteristic functions do not depend on knowing how to perform the following inversion, we include it nevertheless.

Theorem 2.5. *Let $\phi(u)$ be the characteristic function of a distribution function $F(x)$ such that*

$$\int |\phi(u)| du < \infty.$$

Then $F(x)$ has a bounded continuous density $f(x)$ with respect to Lebesgue measure given by

$$f(x) = \frac{1}{2\pi} \int e^{-iux} \phi(u) du.$$

2.4 The Normal Distribution

Definition 2.6. The normal distribution with mean μ and variance σ^2, denoted $N(\mu, \sigma)$, is the distribution of the random variable $\sigma X + \mu$, where

$$P(X < x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt.$$

The normal, or Gaussian, distribution plays a significant role of probability theory because it is the limiting distribution of suitably normalized sums of independent, identically distributed random variables with a finite second moment (variance). The formal statement of this fact is termed the central limit theorem:

Theorem 2.7. *The Central Limit Theorem. Let $X_1, X_2, \ldots$ be independent, identically distributed random variables with $E(X_1) = 0$ and $Var(X_1) = \sigma^2 < \infty$. Then*

$$\frac{X_1 + \cdots + X_n}{\sigma\sqrt{n}} \xrightarrow{\mathcal{D}} N(0,1).$$

2.5 Brownian Motion

As Lawler (1995) notes, Brownian motion is a stochastic process that models random continuous motion, and is an example of a stochastic process with both continuous time and a continuous state space. The terms "Brownian motion" and "Wiener process" are similar, although the second is more general. In particular, a Brownian motion is a Wiener process with a constant variance parameter σ^2. Formally, we define Brownian motion as follows.

Definition 2.8. A Brownian motion, or a Wiener process with variance parameter σ^2 and mean parameter μ, is a stochastic process $X(t)$ taking values in the real numbers satisfying

- *(i)* $X(0) = 0$.
- *(ii)* For any $t_n > t_{n-1} > \cdots > t_0 \geq 0$, the random variables $X_{t_k} - X_{t_{k-1}}$, $k = 1, \ldots, n$ are independent.
- *(iii)* For any $\tau > 0$ and $t \geq 0$, the random variable $X(t+\tau) - X(t)$ has a normal distribution with mean $\mu\tau$ and variance $\tau\sigma^2$.
- *(iv)* The paths are continuous, i.e., the function $t \mapsto X_t$ is a continuous function of t.

Standard Brownain motion is a Brownian motion with $\sigma^2 = 1$. We can also speak of a Brownian motion starting at x; this is a process satisfying conditions (i) through (iv) and the initial condition $X(0) = x$ (Lawler, 1995). Without going into further detail, it is possible to construct Brownian motion as the continuous limit of a random walk in discrete time and discrete state space.

2.6 Poisson Random Variables

Definition 2.9. A random variable X taking values in $\{0, a, 2a, 3a, \ldots\}$ is said to have a Poisson distribution with parameter $\lambda \geq 0$ and jump size a if

$$P(X = ak) = \frac{\lambda^k}{k!}e^{-\lambda}.$$

Equivalently, X has the same distribution if its characteristic function takes the form

$$\phi(u) = \exp\left[\lambda(e^{iua} - 1)\right].$$

2.7 The Poisson Process

The Poisson process is the simplest of the processes with stationary, independent increments. Brownian motion shares these properties with the Poisson process, but unlike Brownian motion, the Poisson process is discontinuous. We say it is the simplest such process because its sample paths are constant except for upward jumps. In particular,

Theorem 2.10. *A process $X(t)$ with stationary, independent increments has a version with all sample paths constant except for upward jumps of length one if and only if there is a parameter $\lambda \geq 0$ such that*

$$E\left(e^{iuX(t)}\right) = e^{\lambda t(e^{iu}-1)}.$$

By expanding, we find that $X(t)$ has the Poisson distribution

$$P(X(t) = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}.$$

2.8 Geometric Brownian Motion

Geometric Brownian motion is the stochastic process used in the Black-Scholes methodology to model the evolution of prices in time. Following Ross (1999), we define geometric Brownian motion as follows. Let the present time be time 0, and let $S(y)$ denote the price of the security at a time y from the present. We say that the collection of prices $S(y), 0 \leq y < \infty$, follow a geometric Brownian motion with drift parameter μ and volatility parameter σ if, for all nonnegative values of y and t, the random variable

$$\frac{S(t+y)}{S(y)}$$

is independent of all prices up to time y; and if, in addition,

$$\log\left(\frac{S(t+y)}{S(y)}\right)$$

is a normal random variable with mean μt and variance $t\sigma^2$. That is, the series of prices will be a geometric Brownian motion if the ratio of the price a time t in the future to the present price will, independent of the past price history, have a lognormal distribution with parameters μt and $t\sigma^2$. The independence assumption is what makes geometric Brownian motion a Markov process.

To put this more directly, we say that $S(t)$ follows a geometric Brownian motion if $\log S(t)$ follows a Brownian motion. In other words, geometric Brownian motion is the exponential of Brownian motion.

If $S(0)$ is the initial price at time 0, then it turns out that the expected value of the price at time t depends on both the mean and variance parameters of the geometric Brownian motion governing the price evolution. In particular, we have

$$E[S(t)] = S(0)e^{t(\mu+\sigma^2/2)}.$$

The expected price grows exponentially at the rate $\mu + \sigma^2/2$.

2.9 Itô's Lemma

As Hull (2000) remarks, the price of a stock option is a function of the underlying stock's price and time. Moreover, the price of any derivative is a function of the stochastic variables underlying the derivative and time (Hull). Itô's Lemma, a result discovered by the mathematician K. Itô in 1951 (Hull), provides a fundamental insight into the behavior of functions of diffusion processes.

Suppose that the value of a variable x follows an Ito process:

$$dx = a(x,t)dt + b(x,t)dz$$

where dz is a standard Brownian motion (Wiener process with $\mu = 0$ and $\sigma^2 = 1$) and a and b are functions of x and t. The variable x has a drift rate of a and a variance rate of b^2. Itô's lemma shows that a function, G, of x and t follows the process

$$dG = \left(\frac{\partial G}{\partial x} a + \frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial x^2} b^2 \right) dt + \frac{\partial G}{\partial x} b dz$$

where dz is the same Wiener process governing the behavior of dx. The notable feature of this result is that G also follows an Itô process. It has a drift rate of

$$\frac{\partial G}{\partial x} a + \frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial x^2} b^2$$

and a variance rate of

$$\left(\frac{\partial G}{\partial x} \right)^2 b^2$$

The notation used above to denote the Wiener process, dz, and the Itô process, dx, requires clarification. The term dz denotes a normal distribution with mean 0 and variance dt. This definition becomes precise by considering discrete time intervals Δt and passing to the limit. In particular, we use the random process $\Delta z = \epsilon\sqrt{\Delta t}$ based on the random variable ϵ, which has mean zero and variance one. Likewise, the process dx is obtained by considering the discrete process just described and passing to the limit as $\Delta t \to 0$.

We will use Itô's lemma without providing a complete proof of the result, which can be viewed as an extension of results in differential calculus to functions of stochastic variables.

3 The Black-Scholes Equation

3.1 Background

The Black-Scholes differential equation must be satisfied by the price, f of any derivative dependent on a non-dividend-paying stock (Hull). The main economic principles used to formulate the differential equation, no-arbitrage and the creation of a riskless portfolio, are simple yet powerful.

The principle of no-arbitrage holds that in a perfectly competitive, liquid market there exist no opportunities to earn a risk-free profit. Stated another way, any portfolio of financial instruments that is perfectly insured against price risk must earn the risk-free interest rate, r.

In order to use the principle of no-arbitrage to value a portfolio, we must first construct a portfolio that is insured against price-risk. The risk that a stock price will change over time will be formulated rigorously in a moment, but the general idea is that at any time, t, the proportion of each instrument in the portfolio must be set so that the net effect of a small change in the price of the underlying asset on the value of the portfolio is zero.

If this condition is satisfied, then it follows that the value of the portfolio at the end of a "short" period of time is known with certainty. Using the principle of no-arbitrage allows us to value a risk-less portfolio over this "short" time interval by setting the percentage return of the portfolio equal to the risk-free interest rate.

3.2 Methodology and Assumptions

The way the no-arbitrage principle is used in the derivation of the Black-Scholes differential equation depends on a couple of assumptions about the market. First, it must be possible to buy or sell any finite quantity of the underlying security at any time. This is what is meant by "perfect liquidity." Second, security trading is continuous in time. In all cases, the full use of the proceeds from buying or short selling securities is permitted. Formally, the set of assumptions used to derive the Black-Scholes equation are as follows (Hull):

Assumptions

1. The stock price follows a geometric Brownian motion with μ and σ constant.
2. The short selling of securities with the full use of proceeds is permitted.
3. There are no transactions costs or taxes.
4. All securities are perfectly divisible.
5. There are no dividends during the life of the derivative.
6. There are no riskless arbitrage opportunities (the no-arbitrage principle).
7. Security trading is continuous.
8. The risk-free rate of interest, r, is constant and the same for all maturities.

Any of these assumptions may be relaxed in order to improve upon the original Black-Scholes model, but some assumptions are more essential than others. While it is relatively easy to incorporate the payment of dividends on the stock into the option pricing methodology, for instance, relaxing the no-arbitrage assumption would be very difficult. This paper is concerned primarily with relaxing the first assumption, and to this end, we present several alternatives to geometric Brownian motion for modeling the price process, as well as several stochastic models for the variance σ^2. First, we present a derivation of the original Black-Scholes equation for a European call option.

3.3 Derivation

The European Call Option: Definition and Discussion

A European call option with strike price K and maturity T gives the buyer the right, but not the obligation, to purchase the underlying stock for $\$K$ on the maturity date (or expiry) T. We will use $f(S,t)$ to denote the price of the option when the underlying stock is trading at price S at time t.

Thus, in the case of the European call option, $f(S,T) = \max(S-K,0)$. This boundary condition follows from the fact that, at time $t = T$, the call option is worth the difference between the stock price and the strike price if the stock price is greater. If the strike price is greater, it is worth nothing at all. In the former instance, the option holder would simply exercise the option to purchase the stock for $\$K$ and simultaneously sell the stock in the market for $\$S$ for a net profit of $\$(S-K)$. This fact will be instrumental in obtaining a closed form solution for the price f of the option as a function of the stock price and time.

The Price Process

Having specified the nature of the derivative contract and determined its boundary condition at time T, the next step is to construct a portfolio consisting of a certain proportion of shares and options that is instantaneously riskless. In order to do this, we must first specify a stochastic process for the stock price. In their landmark paper, Black and Scholes assumed that the stock price follows a geometric Brownian motion:

$$dS = \mu S dt + \sigma S dz \tag{1}$$

This can be rewritten as

$$dS/S = \mu dt + \sigma dz \tag{2}$$

which illustrates that the instantaneous distribution of returns, dS/S, follows a Brownian motion with drift μ and variance σ^2. One attractive feature of geometric Brownian motion is that the stock price S never falls below 0, which accords with the limited liability feature of equity.

It is possible, using Itô's lemma, to consider more general diffusion processes and even some processes with jumps, but we leave that discussion for later.

The Option Price Diffusion

For now, using the assumption that the stock price follows a geometric Brownian motion, and letting $f(S,t)$ denote the price of the option, we invoke Itô's lemma to conclude that the option must follow the diffusion

$$df = \left(\frac{\partial f}{\partial S}\mu S + \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2 \right) dt + \frac{\partial f}{\partial S}\sigma S dz \tag{3}$$

The discrete versions of equations 1 and 3 are

$$\Delta S = \mu S \Delta t + \sigma S \Delta z \tag{4}$$

and

$$\Delta f = \left(\frac{\partial f}{\partial S}\mu S + \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2 \right) \Delta t + \frac{\partial f}{\partial S}\sigma S \Delta z \tag{5}$$

where ΔS and Δf are the changes in f and S over the small time interval Δt.

The Riskless Portfolio

To eliminate the Wiener process Δz, we purchase a portfolio consisting of -1 derivative contracts and $\frac{\partial f}{\partial S}$ shares of the underlying stock. In other words, the holder of this portfolio is short one derivative and long an amount $\frac{\partial f}{\partial S}$ of shares. Define Π as the value of this portfolio. By definition,

$$\Pi = -f + \frac{\partial f}{\partial S} S \tag{6}$$

The change $\Delta\Pi$ in the value of the portfolio over the time interval Δt is given by

$$\Delta\Pi = -\Delta f + \frac{\partial f}{\partial S}\Delta S \tag{7}$$

assuming that $\frac{\partial f}{\partial S}$ is constant over the time period Δt. Substituting equations 4 and 5 into equation 7 and cancelling out the Wiener term yields

$$\Delta\Pi = \left(-\frac{\partial f}{\partial t} - \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2 \right) \Delta t \tag{8}$$

Applying the No-arbitrage Assumption

If all the assumptions we have made are correct, the above equation does not contain a stochastic term and the portfolio must be riskless during the time interval Δt. Therefore, according to the principle of no-arbitrage, the portfolio's return must be equal to the risk-free rate of interest, r. If the return of the portfolio were greater than r, then arbitrageurs could make a risk free profit by shorting risk-free bonds that earn return r and using the proceeds to buy the portfolio. If the portfolio earned a return less than r, arbitrageurs could short the portfolio and buy risk-free bonds. The no-arbitrage argument implies that the percentage return of the portfolio over the time interval Δt should equal r :

$$\Delta\Pi = r\Pi\Delta t$$

By substituting equation 6 for the value of the portfolio and equation 7 for the change in the portfolio into the above equation, we obtain

$$\left(-\frac{\partial f}{\partial t} - \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2\right)\Delta t = r\left(-f + \frac{\partial f}{\partial S}S\right)\Delta t$$

Rearraging yields the famous Black-Scholes differential equation:

$$\frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 f}{\partial S^2} = rf \tag{9}$$

This differential equation has many solutions; obtaining the correct solution for a given financial derivative depends on specifying the correct boundary condition. For the European call option, we showed previously that the correct boundary condition is $f = \max(S-K, 0)$ at time $t = T$.

Obtaining a Closed Form Solution

As Black and Scholes note in their original paper, there is only one formula $f(S,t)$ that satisfies the differential equation 9 subject to the boundary condition $f(S,T) = \max(S-K,0)$. This formula must be the option valuation formula.

To solve the differential equation 9, we make the following substitution:

$$f(S,t) = e^{r(t-T)}g\left[A(S,t), B(S,t)\right]$$

where $A(S,t) = (2/\sigma^2)(r - \frac{1}{2}\sigma^2)[\ln(S/K) - (r - \frac{1}{2}\sigma^2)(t-T)]$

and $B(S,t) = -(2/\sigma^2)(r - \frac{1}{2}\sigma^2)^2(t-T)$

With this substitution, the Black-Scholes differential equation simplifies to

$$\frac{\partial g}{\partial b} = \frac{\partial^2 g}{\partial a^2} \tag{10}$$

and the boundary condition becomes:

$$\begin{aligned} g(a,0) &= K\left[e^{a(\frac{1}{2}\sigma^2)/(r-\frac{1}{2}\sigma^2)} - 1\right] && \text{if } a \geq 0 \\ &= 0 && \text{if } a < 0 \end{aligned}$$

The differential equation 10 is the heat-transfer equation of physics, and its solution is given by Churchill (1963, p.155). By putting the solution into our notation and substituting the resulting function $g(a, b)$ into equation 3.3 and simplifying, we obtain:

$$
\begin{aligned}
f(S,t) &= SN(d_1) - Ke^{r(t-T)}N(d_2) \\
\text{where} \quad d_1 &= \frac{\ln(S/K)+(r+\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} \\
d_2 &= \frac{\ln(S/K)+(r-\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}
\end{aligned}
$$

and $N(x)$ stands for the cumulative probability distribution function for a variable that is normally distributed with a mean of zero and a standard deviation of 1. As specified previously, S is the stock price at the present time ($t = 0$), K is the strike price, r is the continuously compounded risk-free interest rate, σ is the stock price volatility, and T is the time to maturity of the option.

3.4 Properties of the Call Option Formula

We can gain intuition about the call option formula given by the Black-Scholes equation by considering what happens when the parameters take extreme values. For instance, when the stock price S becomes much larger than the strike price K, the probability that the option will be exercised becomes very close to unity. As Hull (2000) notes, the call option becomes very similar to a forward contract (the obligation to pay a certain price at a certain time in the future for delivery of the stock on that date) with delivery price K, which has price

$$S - Ke^{-rT}$$

In fact, as $S \to \infty$ we have $N(d_1) \to 1$ and $N(d_2) \to 1$, so that the call price f approaches the expected price $S - Ke^{-rT}$. Note that if we let the strike price K approach zero for a fixed stock price S, we obtain the same formula. This makes sense, because these parameters appear in the formulas d_1 and d_2 only in the ratio S/K.

When the volatility σ approaches zero, the stock becomes almost riskless. In this case, we expect it to behave much like a risk-free bond, whose price grows at the rate r. Thus, at time T, the payoff of the call option for the (riskless) stock is

$$\max(Se^{rT} - K, 0)$$

The present discounted value of the call today is

$$e^{-rT}\max(Se^{rT}-K,0)=\max(S-Ke^{-rT}),0$$

To show that the Black-Scholes formula is consistent with this prediction, we must consider two cases. First, assume that $S > Ke^{-rT}$. This inequality can be rewritten as $\ln(S/K)+rT > 0$. From the equations for d_1 and d_2, we see that when $\sigma \to 0$, $d_1 \to \infty$ and $d_2 \to \infty$. Thus $N(d_1) \to 1$ and $N(d_2) \to 1$, and we obtain $f = S - Ke^{-rT}$ in the limit. When $S < Ke^{-rT}$, we find that as $\sigma \to 0$, $d_1 \to -\infty$ and $d_2 \to -\infty$. Thus $N(d_1)$ and $N(d_2)$ tend to zero, and in the limit we obtain $f = 0$ for the call price. This confirms that when σ tends to zero, we obtain a call price of $\max(S - Ke^{-rT}, 0)$.

Let us now examine the case when $\sigma \to \infty$. From the formulas for d_1 and d_2, it is easy to determine that in the limit as $\sigma \to \infty$, $d_1 \to \infty$ and $d_2 \to -\infty$. This implies that $N(d_1) \to 1$ and $N(d_2) \to 0$, so in the limit we obtain a call price of $f = S$. In other words, in a world of wildly uncertain stock prices, the strike price is relatively unimportant; to acquire the right to purchase a stock tomorrow for a strike price K, one would just as soon pay the current stock price S.

3.5 Concluding Comment

Having presented a careful derivation of the famous Black-Scholes result and examined the properties of the resulting formula for the price of a European call, we turn a critical eye to one of Black-Scholes' most basic assumptions: the price process.

4 Relaxing the Price Process Assumption

4.1 Motivation

Geometric Brownian motion is the original model for the stock price diffusion on which the Black-Scholes equation is based. While this model is very good as a first approximation for price changes, it can be improved upon substantially. In the past three decades, there have been several theoretical and empirical papers that address this issue. Nevertheless, the price process problem is far from solved: as Robert Almgren noted recently in the January 2002 issue of The American Mathematical Monthly, "constructing improved models for asset price motion and for option pricing is a subject of active research."

Cox and Ross (1976) pointed out 26 years earlier that the critical factor in the original Black-Scholes analysis and in any contingent claims valuation model is the precise description of the stochastic process governing the behavior of the basic asset. In particular, it is "the characteristics of this process that determine the exact nature of the equivalence between packages of claims" (Cox and Ross, 1976). As Merton (1973) concurs, the critical assumptions in the Black-Scholes derivation is that trading takes place continuously in time and that the price dynamics of the stock have a continuous sample path with probability one.

Both Cox and Ross (1976) and Merton (1976) provide a highly useful examination of the option pricing problem for alternative stochastic processes. Other works that treat this issue from an analytical standpoint are Rubinstein (1983), Madan and Seneta (1990), Madan et. al (1998), and Bakshi et al. (1997).

Generally speaking, there are two paths to take for relaxing the assumption that stock prices follow a geometric Brownian motion. First, one may specify an alternative stochastic process for the price and use arguments similar to those used by Black and Scholes (1973) to arrive at the appropriate differential equation, which may be solved using the boundary condition given by the option. There is no guarantee, of course, that the differential equation obtained by this procedure will allow a closed form solution. In difficult cases, numerical procedures may be used along with techniques such as a change of variables to generate solutions for the option price. Alternately, one may use the Fokker-Planck forward equations for the risk-neutral probability functions obtained from the price process to arrive at the appropriate differential equation (as in Bakshi et al., 1997). Although risk-neutral pricing is a useful technique, a lengthy discussion of it is beyond the scope of this paper.

The second path for relaxing the geometric Brownian motion assumption is to specify a stochastic model for the stock price volatility, $\sigma(S, t)$. This method has been used extensively in the literature, and is powerful. A lengthy discussion of it is beyond the scope of this paper, although some of the descriptive empirical models mentioned in the following pages make provisions for non-contant volatility.

Naturally, the paths of changing the price process assumption explicitly and relaxing the assumption of constant volatility are related, and it is possible to accomplish both at the same time. Bakshi et al. (1997) provides a probing empirical study of practically relevant option pricing issues that are related to a number of hybrid stochastic models for volatility and returns.

After discussing the rationale for using geometric Brownian motion in the first place, we review findings in the empirical literature about the actual distribution of returns. We then present several alternative models to geometric Brownian motion that have particular theoretical and analytical strengths. This models include alternative diffusion processes, jump processes, and mixed processes. Technical descriptions of each of the models mentioned in the empirical discussion are included in the latter subsection as well. In the final section of the paper, "Pricing Options for Alternative Stochastic Processes," we discuss the general theory of pricing options for alternative prices processes and apply this methodology to several of the processes mentioned here.

4.2 Why Geometric Brownian Motion?

In section 3.3, we gave the following stochastic process for prices, which was used to derive the Black-Scholes differential equation for options:

$$dS/S = \mu dt + \sigma dz$$

This is the differential form for writing geometric Brownian motion, which is a reasonable "first guess" to use as a model for the price process for several reasons. First, as the above equation indicates, the percentage change in price, or return, equals a drift term (expected return) plus a normally distributed term. This seems more reasonable than the original Brownian motion model for prices proposed by Bachelier (1900),

$$dS = \mu dt + \sigma dz$$

which assumes that the magnitude of price variations is unrelated to the stock price. Intuitively, we would expect the magnitude of price variation to increase with increasing stock price, and geometric Brownian motion takes this feature into account. Also, if the stock has no risk ($\sigma = 0$), then the stock price will grow in time just like a risk-free bond with instantaneous rate of return μ. Naturally, in a world of no-arbitrage, we would have $\mu = r$ in this case. By contrast, for Brownian motion, $\sigma = 0$ implies that $S = \mu t$, which does not accord with the growth formula for a risk-free bond. Finally, geometric Brownian motion is a Markov process with an easily computable law, which makes it very attractive from an analytical point of view.

Empirically, we find that geometric Brownian motion provides a much better fit to the distribution of actual stock price changes than Brownian motion. Also, by using geometric Brownian motion, the form we obtain for the Black-Scholes differential equation can be transformed to the heat equation of physics and solved in closed form. Deriving the Black-Scholes solution using risk-neutral pricing also illustrates the tractability of geometric Brownian motion.

Nevertheless, geometric Brownian motion is not the most accurate model that could be used for the evolution of stock prices. Other diffusion processes besides either Brownian motion or geometric Brownian motion may by used to model stock price evolution. More significantly, the sample paths of stock price evolution may not be continuous, and jump processes can be used to address this discrepancy. On an empirical level, several alternative distributions to simple geometric Brownian motion have been shown to describe the distribution of share price changes with greater accuracy.

4.3 Empirical Evidence on the Return Distribution: An Overview

There have been several papers investigating the statistical qualities of share market returns. Among the most important are: Mandelbrot (1963), Fama (1965), Press (1967), Officer (1972), Praetz (1972), Kon (1984), and Madan and Seneta (1990). Also, there has been a recent trend in the applied physics literature of examining the statistical distibution of prices, in particular the tails. Plerou et al. (1999) is a typical example of such research.

Several of these papers also include new analytical models for stock price changes that are relevant for the option pricing problem. The Madan and Seneta (1990) paper, for instance, presents a Variance Gamma (V.G) model and contains a section on the empirical relevance of the V.G. model for stock market returns. The empirical research in the applied physics literature on stock price time series has led to the proposal of a truncated Levy distribution in order to reproduce the properties of stock prices on different time scales. One of the earliest viable alternatives to the normal distribution for the description of price changes was proposed by Mandelbrot (1963), and in fact, the recently proposed truncated Levy distribution is closely related to this work.

4.4 The Stable Paretian Distribution

In his 1963 paper "The Variation of Certain Speculative Prices," Benoit Mandelbrot pointed out that "the empirical distributions of price changes are usually too 'peaked' to be relative to samples from Gaussian populations." The tails of the distributions of price changes, Mandelbrot claims, are "so extraordinarily long that the second moments typically vary in an erratic fashion." In light of these features, Mandelbrot proposed the stable Paretian distribution as an alternative to the Gaussian. It is important to note that Mandelbrot's initial study examined the distribution of changes in cotton and wool prices, rather than stocks. Nevertheless, the stable Paretian distribution he proposed has been tested as a possible model for the distribuion of stock prices.

The next major work to examine the random nature of price changes, in particular of stocks, was Fama's 1963 paper "The Behavior of Stock-Market Prices." In this piece, Fama concurs with Mandelbrot that the empirical distributions of price changes are leptokurtic (have fat tails), and rejects the hypothesis that price changes can be described by a normal distribution. In particular, his results indicate that the daily changes in log price of stocks of large mature companies follow stable Paretian distributions with characteristic exponents close to 2, but nevertheless less than 2. In other words, Fama claims, "the Mandelbrot hypothesis seems to fit the data better than the Gaussian hypothesis."

The other half of Fama's study was devoted to examining the independence of successive price changes. His conclusion, which falls unambiguously in favor of the random walk hypothesis, is that successive price changes appear to be independent. This supports the hypothesis that the central limit theorem should apply, after some (perhaps lengthy) time scale, to either the price change distribution or the return distribution. However, the distribution of price differences, or log price differences, might not be identical for successive price changes. The distribution of price changes may, for instance, have a volatility that is a function of the stock price–as in geometric Brownian motion. Throughout this paper, we defer to the empirical evidence of Fama and others in support of the independence of successive price change distributions, and concentrate instead on examining the form of the distributions.

4.5 The Compound Events Model

In 1967, Press proposed the "compound events model." The model assumes that log price changes follow a distribution that is a Poisson mixture of normal distributions. Such a distribution, Press comments, is in general skewed, leptokurtic, more peaked at its mean than the distribution of a comparable normal distribution, and has greater probability mass in its tails than a comparable normal distribution. The results of Mandelbrot (1963) and Fama (1963) show that these general properties characterize the distribution of price changes (for both commodities and stocks). However, as Press notes, “there is no need to conclude, on the basis of this evidence, that the variance is infinite, but only that, because of non-zero higher order cumulants, observations will be found further from the mean and that the model ordinate will be higher than would be expected on the basis of normal theory.”

Press tests his compound events model graphically on the price change distributions of several well-known stocks as follows. First, he computes and graphs the cumulative distribution function of actual stock price changes. Second, he estimates the parameters of his model. Finally, he plots the “estimated” theoretical cumulative distribution function on the same graph as the actual c.d.f. by plugging in the estimated parameter values. The agreement between the c.d.f.’s is not always good, as he notes, although the fit is excellent for some stocks. Press attributes the divergence between the theoretical and empirical distributions to a small sample size of stock data. In general, the compound events model of Press possesses the right general properties, but must be calibrated with greater accuracy if it is to provide a workable model for option pricing. In contrast to the stable Paretian distribution, it possesses the significant theoretical advantage of having a finite second moment, which allows the use of standard statistical theory.

4.6 The Statistical Study of Officer (1972)

In 1972, five years after Press published his compound events model for stock returns, R.R. Officer wrote a paper, “The Distribution of Stock Returns,” that took issue with some but not all of the conclusions of Mandelbrot and Fama. Like the previous studies, Officer found that the distribution of stock returns is “fat-tailed” relative to a normal distribution. However, he also observed characteristics that were inconsistent with a stable non-normal generating process. In particular, Officer presents evidence illustrating a tendency for longitudinal sume of daily stock returns to become “thinner-tailed” for larger sums, but not to the extent that a normal distribution approximates the distribution. Also, the standard deviation as a measure of scale appears to be well-behaved.

To test the stability hypothesis for both monthly and daily return data, Officer computes the characteristic exponent α and observes its changes over the sample. For monthly returns, the behavior of α is somewhat inconsistent with both the stable Paretian hypothesis, which predicts a smaller measured α over larger time intervals, and the compound events model of Press, which predicts an increase in the measured α to 2.0, which is characteristic of the normal distribution. In other words, for monthly returns, α appeared to be approximately constant. On balance, this provides evidence that not much is lost by assuming that the distribution of monthly returns is stable.

For daily returns, however, Officer finds a slight increase in the α parameter, which suggests that a "modified model with a finite second moment" for the return distribution might be appropriate. This provides solid support for the compound events model of Press. Further support is provided for the compound events model by the fact that, according to Officer, the standard deviation of the daily sums of returns seems to be well-behaved. The findings on daily returns are probably more important to the portfolio manager, who is likely to rehedge more frequently than once a month, if not daily. Consequently, it is the distribution of daily returns that is important from the perspective of option pricing.

4.7 The t-disribution for Returns

One class of "fat-tailed" distributions with finite second moments and the potential to give a better approximation to the distribution of stock returns is the t-distribution proposed by Praetz (1972). Officer references this paper, which was published earlier in the same year, at the end of his own work. Although the framework provided by Praetz may be used to analyze the distribution of individual share returns or the distribution of the returns on stock indices, Praetz conducts his empirical work solely on stock indices.

Using data from 17 indices on the Australian stock market, Praetz conducted a Chi-squared goodness of fit test for the scaled t-distribution, the normal distribution, the compound events distribution, and the stable Paretian distribution. Using a 1% leve of significance, Praetz concludes, "the results are almost unanimous as all the indices are well-fitted by the scaled t-distribution, whereas the other distributions are rejected in all cases except four." However, he continues, "even in these cases, the scaled t-distribution has a far better fit." For individual share prices, Praetz notes, the situation is not as hopeful due to the discrete nature of the price changes and, in particular, to the large number of zero price changes that always seem to occur. This difficulty cited by Praetz might be remedied by examining data for stocks of large, frequently traded companies. In addition, with the growth of the equities market and of data recording facilities, a much greater volume of data is now available on price changes of individual companies.

4.8 The Discrete Mixture of Normal Distributions

In his 1984 paper "Models of Stock Returns–A Comparison," Stanley Kon proposed a discrete mixture of normal distributions to explain the observed significant kurtosis and "significant positive skewness" in the distribution of daily rates of returns for a sample of common stock and indexes. Stationarity tests on the parameter estimates of this discrete mixture of normal distributions model, Kon claims, revealed significant differences in the mean estimates that can explain the observed skewness and significant differences in the variance estimates that can explain the observed kurtosis. Kon compares the discrete mixture of normal distributions model with the t-distribution model of Praetz, and concludes that the discrete mixture of normal distributions model has "substantially more descriptive validity."

4.9 The Variance Gamma Model for Stock Market Returns

Madan and Seneta (1990) propose a stochastic process called the Variance Gamma (V.G.) model of returns. We will describe the mechanics of the model in more detail later, but essentially, the model stipulates that the unit period distribution is normal conditional on a variance that is distributed as a gamma variate. Empirically, the authors claim, the V.G. model is a good contender for describing daily stock market returns. Madan and Seneta (1990) compare the V.G. model with the normal, the stable Paretian, and the Press compound events model using a Chi-squared goodness-of-fit statistic on seven class intervals for unit sample variance data on 19 stocks quoted on the Sydney stock exchange. For 12 of the 19 stocks studied, minimum chi-squared was attained by the V.G. model. The remaining 7 cases were best characterized by the Press compound events model for five cases, the stable for two cases, and none for the normal distribution.

4.10 Summary of Empirical Findings

Of the models for stock returns and price changes examined in the empirical literature, the ones that best fit the data seem to be the Variance Gamma model of Madan and Seneta (1990) and the Discrete Mixture of Normal Distributions model of Kon (1984). The compound events model of Press (1967) is also a possible contender.

The evidence of slow convergence to a Gaussian behavior over large time scales indicates that the stable Paretian distribution is probably not the best descriptive model for price changes. Nevertheless, modifications of the stable distribution are possible that may provide a better description of the slow convergence to Gaussian behavior. Mantegna and Stanley (1994) have proposed such a process, termed the Truncated Levy Flight, which we will examine later.

Overall, the distribution of stock returns has the following characteristics, which any descriptive model should attempt to take into account: (1) fat-tails and a peaked center relative to the normal distribution; (2) slow convergence to Gaussian behavior for sums of daily returns expressed as log price differences; (3) non-constant variance and mean over time.

Various subsets of the models discussed have been tested against each other using different data sets, but no comprehensive empirical study has been completed so far as we are aware. Ultimately, more empirical work must be done to systematically compare the descriptive performance of the models discussed here on returns for prominent US stocks and US stock indices.

5 Development of Alternative Processes

In presenting candidate processes to replace geometric Brownian motion, we do not strictly follow the order of the processes mentioned in the previous section on empirical findings. Rather, we present the processes first that have been developed the most analytically: alternative diffusions, jump processes, mixed jump diffusion processes, and the variance gamma model. After these, we follow with comparatively brief expositions of the stable Paretian distribution, the compound events model, the t-distribution, the discrete mixture of normal distributions, and the truncated Levy flight.

5.1 Alternative Diffusion Processes

Any Wiener diffusion process may be considered as the limiting case of a general jump process, which we will discuss momentarily. For the time being, consider the general Wiener process

$$dS = \mu(S,t)Sdt + \sigma(S,t)Sdz \tag{11}$$

with drift $\mu(S,t)$ and variance $\sigma^2(S,t)$. The original reasoning of Black and Scholes leads to the following differential equation, which is a generalization of equation 9:

$$\frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2(S,t)\frac{\partial^2 f}{\partial S^2} = rf \tag{12}$$

Note that the above differential equation does not depend on $\mu(S,t)$. The solution obtained from equation 12 does depend explicitly on the form chosen for $\sigma(S,t)$, and we address this problem explicitly in section 5.

The Constant Elasticity of Variance Model

It is useful for the time being to mention a particular class of choices for $\sigma(S,t)$, suggested originally by Cox and Ross and referred to as the "Constant Elasticity of Variance Model" by Hull (2000). In differential notation, the model for the stock price S is

$$dS = \mu S dt + \sigma S^{1-\alpha} dz \tag{13}$$

In this model, $\sigma(S,t) = \sigma S^{1-\alpha}$ for some α with $0 \leq \alpha \leq 1$. Setting $\alpha = 0$ gives geometric Brownian motion.

Two other processes of note, which are discussed by Cox and Ross (1976), are special cases of the Constant Elasticity of Variance Model:

1. Linear Price Variance Process
   $dS = \mu S dt + \sigma \sqrt{S} dz$

2. Constant Price Variance Process
   $dS = \mu S dt + \sigma dz$

The Linear Price Variance Process can be considered as a description of a situation in which changes in state are small and in which the variance of price changes increase with the stock price, but more slowly that in geometric Brownian motion so that the variance of the rate of return decreases rather than remaining constant (Cox and Ross, 1976). Cox and Ross go on to note that, considered in this way, the process can certainly not be rejected on an a priori basis, and may in many situations be preferable to geometric Brownian motion. Unlike geometric Brownian motion, moreover, the Linear Price Variance Process does permit $S = 0$, that is, bankruptcy, to occur with positive probability.

The Constant Price Variance Process would characterize a firm whose price changes have constant variance (independent of the stock price and time). To impose a limited liability constraint, Cox and Ross note, we would let the origin be an absorbing barrier, and consider the equation above as governing the stock price only as long as this point is not reached. With this modification, there would be a positive probability of bankruptcy during any period.

As it turns out, both the Linear Price Variance Process and the Constant Price Variance Process are the diffusion limits of certain jump processes. Cox and Ross (1976) discuss these jump processes in detail.

5.2 Jump Processes

The basic idea behind the use of jump processes to describe the movement of stock prices is that information arrives in chunks, rather than as a continuous stream with no short term surprises. If we let x denote the current state of the world, then the general Markov jump process is of the form

$$dS = \mu(x)dt + dq \tag{14}$$

where the pure jump process is given by

$$dq = \begin{cases} k(x) - 1 & \text{with probability } \lambda(x)dt \\ 0 & \text{with probability } 1 - \lambda(x)dt \end{cases}$$

In the definition of the jump process, $k(x)$ has a distribution dependent on the current world state, x. Note that the above definition is shorthand for a more precise formulation: in particular, in a short time interval Δt, the process will jump zero times with probability $1 - \lambda(x)\Delta t + o(\Delta t)$, once with probability $\lambda(x)\Delta t + o(\Delta t)$, and more than once with probability $o(\Delta t)$. Here the function $o(\Delta t)$ satisfies

$$\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0$$

which ensures that $o(\Delta t)$ is much smaller than Δt for Δt small. As in Cox and Ross (1976), we will assume that $x = S$ to indicate the belief that all state information is contained in the current stock value, S. A Wiener diffusion term, $\sigma(x)dz$, could be added to equation 5.2 to obtain a more general form, but we refrain from doing so because the Wiener diffusion can be arrived at by taking the limit of the jump process.

We will now describe the particular jump processes, which are are special cases of equation 5.2, that in the limit become the Linear Price Variance Process and the Constant Price Variance Process, respectively.

First, let the intensity $\lambda(S) = \lambda S$ and let the drift $\mu(S) = \mu S$. This encodes the assumption that the drift in the rate of return dS/S is constant and the assumption that information tends to arrive more frequently when the stock price is higher. Additionally, chose the distribution for the jump in prices to be independent of price. Thus the distribution for the jump component is given by

$$dq = \begin{cases} k - 1 & \text{with probability } \lambda S dt \\ 0 & \text{with probability } 1 - \lambda S dt \end{cases}$$

so that we have the stochastic process

$$dS = \mu S dt + dq \tag{15}$$

for price changes. As Cox and Ross (1976) note, equation 15 is a generalization of a class of stochastic processes known as birth and death processes. The local mean and variance of 15 are given by

$$E[dS] = (\mu + \lambda E[k-1])Sdt$$

and

$$Var[dS] = \lambda E[(k-1)^2]Sdt$$

To construct a pure birth and death process we ignore the drift in equation 15 and let the random variable k take on two values, $k^+ > 1$ and $k^- < 1$ with respective conditional probabilities π^+ and π^-. This gives us the stochastic process

$$dS = \begin{cases} k^+ - 1 & \text{with probability } \pi^+ \lambda S dt \\ k^- - 1 & \text{with probability } \pi^- \lambda S dt \\ 0 & \text{with probability } 1 - \lambda S dt \end{cases}$$

Equation 5.2 is an example of a simple birth and death process for a population (Cox and Ross, 1976). Following Cox and Ross (1976), imagine a firm made up of individual units whose sum value (population size) is S. If these units are stochastically independent of each other, we can let λdt represent the probability of an event occurring for any one unit. An event is, with probability π^+, the 'birth' of $k^+ - 1$ additional units and with probability π^- the 'death' of $1 - k^-$ units. For the whole firm, then, equation 5.2 describes its local movement.

By taking the limit of equation 5.2 as $k^+ \to 1$ and $k^- \to 1$ and $\lambda \to \infty$ in a particular manner, we obtain the Linear Price Variance Process,

$$dS = \mu S dt + \sigma \sqrt{S} dz$$

Note that in the above equation, μ is not the same as the drift in equation 15. Instead, the μ and σ in the Linear Price Variance equation are given by

$$\mu = \lambda E[k-1]$$

and

$$\sigma = \sqrt{\lambda E[(k-1)^2]}$$

As noted in the previous section, this process does allow the stock price to reach zero with positive probability.

Another specialization of the general Markov jump process 5.2 can be used to obtain the Constant Price Variance Process. To accomplish this, suppose the firm is composed of dependent units, so that the intensity λ is constant, and let the value increment also be constant as in the last process. Then we have the stochastic process

$$dS = \mu S dt + dq,$$

with jump component

$$dq = \begin{cases} k^+ - 1 & \text{with probability } \pi^+ \lambda dt \\ k^- - 1 & \text{with probability } \pi^- \lambda dt \\ 0 & \text{with probability } 1 - \lambda dt \end{cases}$$

Cox and Ross (1976) call this the absolute process. This is a case where value grows endogenously at the exponential rate μ and where lump exogenous changes to value of size $k - 1$ occur with intensity λ.

The local mean and variance of the absolute process are given by

$$E[dS] = (\mu S + \lambda[\pi^+(k^+ - 1) + \pi^-(k^- - 1)])dt$$

and

$$Var[dS] = \lambda[\pi^+(k^+ - 1)^2 + \pi^-(k^- - 1)^2]dt$$

If $\pi^- = 0$ then the process has limited liability, but if $\pi^- > 0$ there is a positive probability that the stock price will reach zero. To preserve limited liability, Cox and Ross (1976) note, we would have to specify a non-negative lower barrier for S and treat $S = 0$ as an absorbing boundary. Taking the diffusion limit of the absolute process 5.2, we obtain the diffusion

$$dS = \mu S dt + \bar{\sigma} dz$$

which is the Constant Price Variance Process. The drift μ is the same as the drift in the absolute process, and the standard deviation σ is given by

$$\sigma = \sqrt{\lambda[\pi^+(k^+ - 1)^2 + \pi^-(k^- - 1)^2]}$$

There is a subtlety here: in taking the limit of the jump process to obtain the diffusion, we do so in a way that maintains the instantaneous mean and variance, and we set the mean of the jump process to zero so that the resulting Wiener process dz has mean zero. As with the absolute process, the Constant Price Variance Process has a positive probability of bankruptcy in any period, and to impose limited liability we would set the origin as an absorbing boundary.

A natural question to ask at this point is, "which jump process has geometric Brownian motion as its limit?" The answer is the following process, which is similar to the processes discussed above:

$$dS/S = \mu dt + dq \tag{16}$$

where the pure jump component is given by

$$dq = \begin{cases} k-1 & \text{with probability } \lambda dt \\ 0 & \text{with probability } 1-\lambda dt \end{cases}$$

In geometric Brownian motion, the variance of the price changes is proportional to a constant (σ^2) times the square of the stock price. We obtain this limiting behavior because in the above process, the units of value comprising S are completely dependent. This is reflected by the fact that the intensity λ is independent of the stock price S. Thus when a new chunk of information arrive, it affects each unit of the stock price equally. The dependence of events is also characteristic of the absolute process. The difference between the jump processes, which leads to the significant difference between the limiting diffusions, is fact that in the jump process leading to geometric Brownian motion, the jump size is proportional to the stock price S.

To conclude, we note a useful formula by Merton (1976) for the price of a stock that undergoes jumps. Merton considers the stochastic process with a Wiener component as well as a jump component, written

$$dS/S = (\mu - \lambda k)dt + \sigma dz + dq \tag{17}$$

where the pure jump process is given by

$$dq = \begin{cases} Y-1 & \text{with probability } \lambda dt \\ 0 & \text{with probability } 1-\lambda dt \end{cases}$$

In the above equation for dS/S, μ is the instantaneous expected return on the stock; σ^2 is the instantaneous variance of the return conditional on the Poisson event not having occured; dz is a standard Wiener process; dq and dz are assumed to be independent; λ is the mean number of arrivals per unit time (the intensity); $k = E[Y-1]$ is defined as the expected percentage change in the stock price if the Poisson event occurs; and E[] is the expectation operator over the random variable Y. Note that $Y-1$ is an impulse function producing a finite jump in S to SY. The resulting sample path for

$S(t)$, Merton notes, will be continuous most of the time with finite jumps of differing signs and amplitudes occuring at discrete points in time. If μ, λ, k, and σ are constants then the random variable ratio of the stock price at time t to the stock price at time zero (conditional on $S(0) = S$ can be written as

$$S(t)/S = \exp[(\mu - \sigma^2/2 - \lambda k)t + \sigma z(t)]Y(n) \tag{18}$$

where $z(t)$ is a Gaussian random variable with a zero mean and variance equal to t; $Y(n) = 1$ if $n = 0$; $Y(n) = \prod_{j=1}^{n} Y_j$ for $n \geq 1$ where the Y_j are independently and identically distributed and n is Poisson distributed with parameter λt. Equation 18 gives a useful picture of the dynamics of the stock price evolution over time.

5.3 The Variance Gamma Process

One of the more recent models for stock price dynamics that has been proposed is the Variance Gamma (V.G.) model (Madan and Seneta, 1990). The V.G. model satisfies several practical and empirically relevant properties:

1. Long tailedness relative to the normal for daily returns, with returns over longer periods approaching normality (Fama, 1965).
2. Finite moments for at least the lower powers of returns.
3. Consistency with an underlying, continuous-time stochastic process, with independent, stationary increments, and with the distribution of any increment belonging to the same simple family of distributions irrespective of the length of time to which the increment corresponds (thereby permitting sampling and analysis through time in a straightforward fashion).
4. Extension to multivariate processes with elliptical multivariate distributions that thereby maintain validity of the capital asset pricing model (Owen and Rabinovitch, 1983).

We have discussed already the Brownian motion and geometric Brownian motion models for stock price dynamics, as well as several jump processes that can be used in conjunction with geometric Brownian motion (or any other diffusion) or on their own. The literature on market returns contains other models besides these and the one currently under discussion. These models include the symmetric stable distribution proposed by Mandelbrot (1963), the compound events model combining normally distributed jumps at Poisson jump times proposed by Press (1972), the t-distribution suggested by Praetz, and the generalized Beta distribution suggested by Bookstaber and McDonald (1987).

As Madan and Seneta (1990) discuss, Brownian motion fails property 1. The symmetric stable fails on properties 2 and 3. The Praetz t-distribution fails on property 3 because it is not possible to construct a stochastic process with the property 3 and distributions of any increment being a t-distribution irrespective of length of time interval considered since the sum of independent t-variables is not a t-variable.

Though the compound events model of Press possesses all the four properties described above, the authors claims, the proposed V.G. model has a further advantage in being a pure jump process of, in the main, a large number of small jumps. In fact, Madan and Seneta (1990) show that the V.G. model is a limit of a particular sequence of compound events models in which the arrival rate of jumps approaches infinity, while the magnitudes of the jumps are progressively concentrated near the origin.

The Variance Gamma model can be stated formally as follows. Let $R(t)$ be the return over a unit time period. That is, $R(t) = S(t+1)/S(t)$, where $S(t)$ is the stock price at time t. Suppose that $\ln(R(t))$ is normally distributed with mean μ and a random variance $\sigma^2 V$, where μ and σ^2 are known constants. The distribution of V is taken to be gamma, with parameters c, γ, and density $g(\nu)$ given by

$$g(\nu) = \frac{c^\gamma \nu^{\gamma-1} e^{-c\nu}}{\Gamma(\gamma)}, \tag{19}$$

where Γ is the gamma function. If $X = \ln(R) - \mu$ (dropping the t dependence for notational convenience), then the density of X, $f(x)$ is given by

$$f(x) = \int_0^\infty \frac{e^{-x^2/(2\sigma^2\nu)}}{\sigma\sqrt{2\pi\nu}} g(\nu) d\nu \tag{20}$$

which has no closed-form expression. However, the characteristic function for X, $\phi_X(u)$, has a closed-form expression obtained easily by conditioning on V,

$$\phi_X(u) = [1 + (\sigma^2 v/m)(u^2/2)]^{-m^2/v}, \tag{21}$$

where $m = \gamma/c$ is the mean of the gamma density $g(\nu)$ and $v = \gamma/c^2$ is its variance. Since σ^2 serves as the scale parameter for the variance, we can take the mean of V to be $m = 1$.

Madan and Seneta (1990) show that the variable V can be viewed as a random time change and this setting of $m = 1$ is consistent with supposing that the expected random time change is unity for the unit period return. The characteristic function of the unit return distribution therefore is

$$\phi_X(u) = [1 + \sigma^2 v u^2/2]^{-1/v} \tag{22}$$

The continuous-time stochastic process $Y(t)$, which is consistent with the V.G. model as the distribution for the unit period motion $Y(t+1) - Y(t)$, is given by Brownian motion applied to random time change:

$$Y(t) = b(G(t)), \tag{23}$$

where $G(t)$ is the process of i.i.d. gamma increments with mean τ and variance $v\tau$ over intervals of length τ, and $b(t)$ is an independent Brownian motion of zero drift and variance rate σ^2.

5.4 The Stable Paretian Distribution

Mandelbrot (1963) presents several properties of the Levy stable distribution discovered by Paul Levy, and suggests this distribution as a potential model for returns. The first noteworthy fact about the stable distributions is that they are the general class of distributions that satisfy the relation

$$a_1 U + a_2 U \overset{\mathcal{D}}{=} aU$$

where U represents a random variable, a_1 and a_2 represent scale factors, and a is a function of a_1 and a_2. The normal distribution satisfies the above relation, as does the well-known Cauchy distribution. The normal distribution is the only stable distribution with a finite variance. The Cauchy distribution, on the other hand, has an infinite second moment but a finite expectation. In fact, all stable distributions with a finite expectation can be thought of as lying on a continuum between the Cauchy distribution and the Gaussian.

The general characteristic function for the stable distribution is given by the following theorem:

Theorem 5.1. $\phi(u) = e^{\psi(u)}$ *is the characteristic function of a stable law of exponent* $\alpha, 0 < \alpha < 1,$ *and* $1 < \alpha < 2$ *if and only if it has the form*

$$\psi(u) = iuc - d|u|^\alpha (1 + i\theta \frac{u}{|u|} tan(\frac{\pi}{2}\alpha)),$$

where c is real, d is real and positive, and θ *is real with* $|\theta| \leq 1$. *For* $\alpha = 1$, *the form of the characteristic function is given by*

$$\psi(u) = iuc - d|u|(1 + i\theta \frac{u}{|u|} \frac{2}{\pi} \ln(|u|)),$$

where c, d, and θ are as above.

The parameters c, d, and θ are unique, and have useful interpretations. Let us use the notation $X \stackrel{\mathcal{D}}{=} S_\alpha(\sigma, \theta, c)$ to mean that X is a stable random variable with exponent α and parameters $d = \sigma^\alpha, \theta$, and c. The parameter c is a shift parameter, which can be seen from the fact that if $X \stackrel{\mathcal{D}}{=} S_\alpha(\sigma, \theta, c)$, then $X + a \stackrel{\mathcal{D}}{=} S_\alpha(\sigma, \theta, c + a)$. To explain the parameter σ, we write $\sigma = d^{1/\alpha}$ for a stable distribution with exponent α. Then σ is said to be a scale parameter because

$$\begin{array}{ll} kX \stackrel{\mathcal{D}}{=} S_\alpha(|k|\sigma, sign(k)\theta, kc) & \text{if } \alpha \neq 1 \\ kX \stackrel{\mathcal{D}}{=} S_1(|k|\sigma, sign(k)\theta, kc - \frac{2}{\pi}k(\ln|k|)\sigma\theta) & \text{if } \alpha = 1. \end{array}$$

Lastly, θ is said to be a skewness paramater, because $X \stackrel{\mathcal{D}}{=} S_\alpha(\sigma, \theta, c)$ is symmetric if and only if $\theta = 0$ and $c = 0$. Moreover, X is symmetric about c if and only if $\theta = 0$.

Mandelbrot originally proposed the stable distribution to model log price changes of the form

$$L(t, 1) = \ln(S(t + 1)) - \ln(S(t))$$

for a unit time difference. His initial motivation for this choice was the fact that, if the second moment of log price changes diverges but the first moment is well behaved, then the density (call it $p(u)$ for the moment) must decrease faster than u^{-2} but slower than u^{-3}. This requirement is significant because all stable distributions follow the law of Pareto, which states that, for a stable random variable X

$$P(X > q) \sim Kq^{-\alpha},$$

where K is a constant determined by the parameters of the distribution X. A similar relationship holds for the lower probability tail. This relationship is why Mandelbrot uses the term “stable Paretian” to describe Levy stable distributions in the context of price changes. For reference, the Gaussian corresponds to the case $\alpha = 2$ and the Cauchy distribution corresponds to the case $\alpha = 1$ and $\beta = 0$. The applied physics literature refers to this phenomenon as “power law scaling” (see, for ex., Gopikrishnan et al., 1998). The Paretian feature of stable distributions provides a very direct basis for empirical testing. Gopikrishnan et al. (1998), for instance, claim that the distribution of stock price changes obeys an approximately inverse cubic law, which is outside the Levy stable regime ($0 < \alpha < 2$).

5.5 The Compound Events Model

The compound events model of Press (1967) involves a Poisson process with Gaussian jump sizes. To be precise, The compound events model states that

$$\ln(S(t)) - \ln(S(0)) = \sum_{k=1}^{N(t)} Y_k + X(t) \tag{24}$$

where $S(t)$ is a process of stationary, independent increments; $S(0)$ is assumed known; $(Y_1, Y_2, \dots, Y_k, \dots)$ is a sequence of mutually independent random variables each drawn from a normal distribution with mean θ and variance σ_2^2; $N(t)$ is a Poisson process with parameter λt, which represents the number of event occuring in time t; $\{N(t), t \geq 0\}$ is independent of the Y_k; and $\{X(t), t \geq 0\}$ is a Wiener process independent of $N(t)$ and of $(Y_1, Y_2, \dots)$, and $X(t)$ is drawn from a normal distribution with mean 0 and variance $\sigma_1^2 t$.

Let $Z(t) = \ln(S(t)) - \ln(S(0))$ and let $\phi(u) = E[e^{iuZ(t)}]$ be the characteristic function of $Z(t)$. Then, following Press (1967), the characteristic function is given by

$$\ln \phi(u) = iCu - \frac{t\sigma_1^2 u^2}{2} + \lambda t[e^{i\theta u - (\sigma_2^2 u^2/2)} - 1]$$

Now, let $\phi^*(u)$ denote the characteristic function of the one-step price change $\Delta Z(t) = Z(t) - Z(t-1)$. The characteristic function of $\Delta Z(t)$ has the form

$$\ln \phi^*(u) = -\frac{\sigma_1^2 u^2}{2} + \lambda[e^{i\theta u - (\sigma_2^2 u^2/2)} - 1]$$

As Press explains, the degree to which $\phi^*(u)$ departs from normality depends on the magnitude of the coefficients of the terms in the series expansion of the above formula that are higher than quadratic in u. Thus, for λ very small, $\phi^*(u)$ is approximately normally distributed. For λ large, however, the higher order terms produce a substantial departure from normality (Press, 1967). For comparison to some of the previously mentioned jump processes, the means and variances for $\Delta Z(t)$ are

$$E[\Delta Z(t)] = \theta\lambda$$

and

$$Var[\Delta Z(t)] = \sigma_1^2 + \lambda(\theta^2 + \sigma_2^2)$$

It can be shown analytically that the distribution of $\Delta Z(t)$ is leptokurtic and that, when θ is small, the probability in the extreme tails of the distribution of $\Delta Z(t)$ exceeds that of a comparably normally distributed random variable.

5.6 The t-distribution

Praetz (1972) begins by defining the variable of log share price changes $Y = \ln(P(t+\tau)) - \ln(P(t))$. The work of Osborne (1959), which he references, shows that prices can be interpreted as an ensemble of decisions in statistical equilibrium, with properties resembling an ensemble of particles in statistical mechanics. The equilibrium distribution of Y is given by

$$f(y) = \frac{e^{-y^2/(2\sigma^2\tau)}}{\sqrt{2\pi\sigma^2\tau}}$$

where σ^2 is the variance of Y over unit time intervals. This distribution is the same as that of a particle in Brownian motion, and thus the price $P(t)$ is shown to follow a geometric Brownian motion in this case.

The contribution of Praetz is the proposal of a distribution for the volatility parameter σ^2. This is one of the earliest works that proposes a stochastic volatility framework, although the potential significance for option pricing was not clear at the time, since Black and Scholes has not yet published their landmark 1973 paper.

In particular, Praetz proposed that Osborne's formula for the p.d.f. of Y be conditioned on the value of σ^2, that a unit time interval be considered for simplicity, and that Y should have a non-zero drift μ. This changes Osborne's formula to

$$f(y|\sigma^2) = \frac{e^{-(y-\mu)^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}}$$

If we denote by $h(y)$ the distribution of Y that takes into account the distribution of σ^2, then $h(y)$ can be computed by

$$h(y) = \int_0^\infty f(y|\sigma^2)g(\sigma^2)d\sigma^2$$

with $0 \le \sigma < \infty$. An acceptable solution for the distribution of σ^2, Praetz claims, is

$$g(\sigma^2) = \bar{\sigma}^{2m}(m-1)^m\sigma^{-2(m+1)}e^{-(m-1)\bar{\sigma}^2/\sigma^2}/\Gamma(m)$$

Here, $\bar{\sigma}^2 = E[\sigma^2]$ and the variance of σ^2 is $\bar{\sigma}^4/(m-2)$. This is known as an inverted gamma distribution. When $g(\sigma^2)$ is substituted in the equation for $h(y)$ we obtain $h(y)$ by integration as

$$h(y) = [1 + (y-\mu)^2/\bar{\sigma}^2(2m-2)]^{-m-1/2}\Gamma(m)[(2m-2)\pi]^{1/2}\bar{\sigma} \qquad (25)$$

This is a t-distribution of $2m = n$ degrees of freedom, except for a scale factor $[n/(n-2)]^{\frac{1}{2}}$. The distribution of $(Y - \mu)/\bar{\sigma}$, therefore, would be that of a scaled t-distribution. Significantly, we can obtain the distribution of $P(t)$, the price of a share at time t, from $Y = \ln(P(t)) - \ln(P(0))$. We will not reproduce the formula here, but it can be found in Praetz (1972).

5.7 The Discrete Mixture of Normal Distributions

Kon (1984) argues that the true distribution of stock returns may be normal, but its paramaters shift among a finite set of values. There are time-ordered shifts associated with capital structure changes, acquisitions, stock splits, or exogenous market events. There are also cyclical shifts between sets of parameters, as in the day of the week effect or the seasonal announcements of firm earnings and dividends. A model specification intended to represent the true mixture process, therefore, must be able to accomodate both cyclical and structural (time-ordered) shifts in the two parameters of a normal distribution.

The model proposed by Kon assumes that each return observation is a drawing from one of N sets of parameter values. As long as we refer to subsets of the data whose observations are not necessarily consecutive in time, Kon claims, then both the structural and cyclical type parameter shifts can be accomodated.

The generalized discrete mixture of normal distributions model views each return observation on a stock, $r(t)$, as having been generated by one of the following N distinct equations:

$$\begin{aligned} r(t) &= \mu_1 + U_1 && \text{if } t \in I_1 \\ r(t) &= \mu_2 + U_2 && \text{if } t \in I_2 \\ &\vdots && \vdots \\ r(t) &= \mu_N + U_N && \text{if } t \in I_N \end{aligned}$$

where $I_i, i = 1, 2, \ldots, N$ are the homogeneous information sets with T_i observations in each set. Thus, $\sum_{t=1}^{N} T_i = T$. The random variables U_i are independent and each normally distributed with a mean of zero and variances of $\sigma_i^2, i = 1, 2, \ldots, N$ respectively.

Define $\lambda_i = T_i/T$ as the proportion of observations associated with information set I_i. Then, for a given N, the parameter vector $\theta = \{\mu_1, \mu_2, \cdots, \mu_N, \sigma_1^2, \sigma_2^2, \cdots, \sigma_N^2, \lambda_1, \lambda_2, \cdots, \lambda_N\}$ can be estimated by maximizing the likelihood function

$$l(\theta|r) = \prod_{t=1}^{T} \left[\sum_{i=1}^{N} \lambda_i p(r(t)|\gamma_i) \right]$$

where $r = (r_1, r_2, \cdots, r_T)'$, $\gamma_i = (\mu_i, \sigma_i^2)$, and $p(r(t)|\gamma_i)$ is a normal probability density function with mean μ_i and variance σ_i^2. The details of the estimation procedure and the generality of model specification are discussed in the Appendix of Kon's paper.

5.8 The Truncated Levy Flight

Somewhat recently, Mantegna and Stanley (1994) proposed a stochastic process known as the truncated Levy flight, which exhibits a slow convergence to Gaussian behavior. Given the empirical evidence presented up to this point, this process clearly has applications to stock price time series. The truncated Levy flight is constructed as follows. First, consider the sum

$$Z_n = \sum_{i=1}^{n} X_i$$

of n i.i.d. random variables X_i with finite variance. The central limit theorem states that as $n \to \infty$, the random variable Z_n will converge to a normal distribution. The rate of convergence to the normal distribution, however, is not specified by the central limit theorem, and depends on the distribution of the X_i. Suppose each of the random variables X_i shares the same distribution as the random variable X. Then the Truncated Levy Flight (TLF) is characterized by the following probability distribution for X :

$$T(x) = \begin{cases} 0 & \text{if } x > l \\ c_1 L(x) & \text{if } -l \leq x \leq l \\ 0 & \text{if } x < -l \end{cases}$$

where

$$L(x) = \frac{1}{\pi} \int_0^\infty \exp(-\gamma q^\alpha) cos(qx) dq$$

is the symmetrical Levy stable distribution of index α ($0 < \alpha \leq 2$) and scale factor $\gamma > 0$, c_1 is a normalizing constant, and l is the cutoff length. For the sake of simplicity, Mantegna and Stanley (1994) set $\gamma = 1$. While their paper presents a method for quantifying the time until the TLF converges to Gaussian behavior, we omit this discussion here. It suffices to say that, for appropriate values of the parameters for $L(x)$ and the cutoff length l, the TLF may provide a reasonable empirical fit to daily log price data.

6 Pricing Options for Alternative Stochastic Processes

We begin by offering a few comments on option valuation theory, and follow with a concise exposition of option pricing results for alternative diffusion processes, jump processes, and mixed processes. We conlude this section with suggestions for future research on pricing options for processes adapted from the empirical literature.

6.1 A Few Comments on Option Valuation Theory

There are several approaches to option valuation theory, but one of the best frameworks is presented by Cox and Ross (1976), who illustrate the structure of hedging arguments to obtain valuation formulas for options in a fairly general setting. The procedure can be summarized as follows:

1. First, choose a particular stochastic process to govern the price movement of the underlying asset, say a stock with price S.
2. Next, take an instrument whose value is dependent on S, say an option written on the stock, and assume that a sufficiently regular price function $f(S, t)$ exists.
3. Given that the price process and the option price function $f(S, t)$ are sufficiently well behaved, derive the process for the differential movement in the option value, df.
4. Keep in mind that the drift and variance parameters of the option price drift df depend on the unknown function $f(S, t)$ and the known values of S and t.
5. Assume the existence of short-selling and a third riskless asset that earns an instantaneous interest rate r. Assume no-arbitrage.
6. For Poisson stock price processes, Cox and Ross assume the jump amplitude is a non-random function at a jump to ensure tractability.
7. Use the fundamental option valuation equation given in Cox and Ross (1976) to forumlate a differential-difference equation for the option price.
8. Use the terms of the option to set boundary conditions for this differential-difference equation and apply known techniques to solve it.

To flesh out the above framework, we need to mention a couple of things. The random differential movement of S is written as

$$dS = \mu(S,t)dt + \sigma(S,t)dx$$

where $\mu(S,t)$ and $\sigma(S,t)$ are taken to be function of the current state of the world, which for simplicity is supposed to be summarized by S and t alone. The option process df can be expressed as

$$df = \mu(f,t)dt + \sigma(f,t)dx$$

Cox and Ross use a few equations to obtain their fundamental formula for option valuation. First, they use the existence of a hedge portfolio of the stock, S, and the option, $f(S,t)$, to write down the relationship

$$\alpha_S \sigma(S)(dx/S) + \alpha_f \sigma(f)(dx/f) = 0$$

where the dependence on t is dropped for simplicity. This simplifies to

$$\alpha_S(\sigma(S)/S) + \alpha_f(\sigma(f)/f) = 0$$

where α_S and α_f are the portfolio weights in the stock and the option, respectively. Next, since the hedge portfolio is riskless, it must have a rate of return

$$\alpha_S(\mu(S)/S) + \alpha_f(\mu(f)/f) = (\alpha_S + \alpha_f)r$$

From these two equations, we obtain the fundamental option valuation equation

$$\frac{(\mu_f - rf)}{\sigma_f} = \frac{(\mu_S - rS)}{\sigma_S}$$

which states simply that the risk premium divided by the scale of risk has to be the same for the stock and the option. To clarify, we derived the equations leading to the fundamental option valuation equation by seperating out the risky and riskless components of the the essential relation

$$\alpha_S\left(\frac{dS}{S}\right) + \alpha_f\left(\frac{df}{f}\right) = (\alpha_S + \alpha_f)r$$

which encodes the idea that the total return for the stock holdings plus the total return from the option holdings must equal the risk-free return on a sum equal to the dollar amount of stock holdings plus the dollar amount of option holdings.

We will not delve into the details here, but it can be argued that the uniqueness of the solution $f(S,t)$ and the independence of the hedging argument from any presumption about the risk-preferences of investors implies that that risk-neutrality can be assumed in computing the option price function. This implies that the expected return on the stock and the option can be set equal to the risk free interest rate r. The option price, then, would be computed by discounting the expected value of the terminal option price distribution. Sometimes, this procedure yields solutions more easily than the procedure of formulating the differential-difference equation and solving it via standard transformation and series techniques. However, the two approaches can be shown to be equivalent.

6.2 Alternative Diffusion Results

The original argument of Black-Scholes yields the following differential equation for the option price, which was stated earlier:

$$\frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2(S,t)\frac{\partial^2 f}{\partial S^2} = rf \tag{26}$$

For the Linear Price Variance Process, this differential equation becomes

$$\frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S\frac{\partial^2 f}{\partial S^2} = rf \tag{27}$$

Because the density of the limiting diffusion for birth and death processes in known (Feller, 1951), we can apply the risk-neutral technique and take the expectation of $\max(S(T)-K,0)$ discounted to time t to obtain the valuation formula

$$f(S,t) = S\sum_{n=0}^{\infty}\frac{(n+1)e^{-y}y^n G(n+2,\theta K)}{\Gamma[n+2]} - Ke^{-r(T-t)}\sum_{n=0}^{\infty}\frac{e^{-y}y^{n+1}G(n+1,\theta K)}{\Gamma[n+2]}$$

where

$$\theta = \frac{2r}{\sigma^2[e^{r(T-t)}-1]}$$

$$y = \theta S e^{(T-t)}$$

$$G(m,x) = [\Gamma(m)]^{-1}\int_x^{\infty} e^{-z}z^{m-1}dz$$

The value of an option at $S=0$ is implied by the description of the process and no additional restrictions need to be made (Cox and Ross, 1976). For a process with an absorbing barrier at zero we will have $f(0,t)=0$. Note again that the above formula could have been obtained by using standard series techniques on the differential equation.

We turn to the Constant Price Variance Process. The differential equation for the option price becomes

$$\frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial S^2} = rf \tag{28}$$

Here, we impose an absorbing barrier at zero. As with the previous process, the density of the diffusion is known, so (either by transformation to the heat equation or by risk-neutral expectation) we obtain the valuation formula

$$\begin{aligned} f(S,t) = \quad & (S - Ke^{-r(T-t)})N(y_1) \\ & +(S + Ke^{-r(T-t)})N(y_2) \\ & +v[n(y_1) - n(y_2)] \end{aligned}$$

where $N()$ is the cumulative unit normal distribution function, $n()$ is the unit normal density function, and

$$\begin{aligned} v &= \sigma\left(\tfrac{1-e^{-2r(T-t)}}{2r}\right)^{1/2} \\ y_1 &= \tfrac{S-Ke^{-r(T-t)}}{v} \\ y_2 &= \tfrac{-S-Ke^{-r(T-t)}}{v} \end{aligned}$$

We turn now to the valuation of options on jump processes.

6.3 Results for Jump Processes

As throughout this paper, our problem is to value a European call option with an expiration date T on which the holder receives $\max(S(T) - K, 0)$. For a stock whose local return is given by equation 5.2, the option follows the perfectly dependent process

$$df = \begin{cases} f(S+k-1,t) - f(S,t) & \text{with probability } \lambda S dt \\ \frac{\partial f}{\partial t}dt + \mu\frac{\partial f}{\partial S}dt & \text{with probability } 1 - \lambda S dt \end{cases}$$

where λ is an arbitrary function. By constructing the fundamental option valuation equation for this option process, Cox and Ross (1976) obtain the following difference-differential equation:

$$\mu\frac{\partial f}{\partial S} + \left[\frac{\mu - rS}{1-k}\right]f(S+k-1,t) + \left[\frac{r[k-1+S]-\mu}{1-k}\right]f(S,t) + \frac{\partial f}{\partial t} = 0$$

where μ and k are functions of S and t. Note that the above equation is independent of the process intensity λ. When the hedge position depends only on the jump size, in fact, the intensity plays no role in the valuation. The above equation can be used to examine a variety of jump processes. We will not go into the details here, but the formula for pricing a pure birth process without drift is contained in Cox and Ross (1976). In addition, Cox and Ross examine the pricing problem for a pure jump process augmented by a proportional drift term μS. Unfortunately, this problem has no closed form solution.

6.4 Merton's Mixed Process

Merton (1976), like Cox and Ross (1976), examines a jump process with a drift term. Although he does not find a closed form solution either, his formulation of the problem is a bit more compact and his underlying stochastic process is stated in terms of the return dS/S rather that the price increment dS. To review, the return process used by Merton is

$$dS/S = (\mu - \lambda k)dt + \sigma dz + dq \tag{29}$$

where the pure jump process is given by

$$dq = \begin{cases} Y - 1 & \text{with probability } \lambda dt \\ 0 & \text{with probability } 1 - \lambda dt \end{cases}$$

Using his variation of the fundamental option pricing equation, Merton arrives at the following difference-differential equation for the option price $f(S,t)$:

$$\frac{\partial f}{\partial t} + (r - \lambda k)S\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} + \lambda\epsilon\{f(SY,t) - f(S,t)\} = rf \tag{30}$$

subject to the boundary conditions

$$f(0,t) = 0$$

and

$$f(S,T) = \max(S - K, 0)$$

While a complete closed form solution to equation 30 cannot be written down without a further specification for the distribution of Y, a partial solution that is in reasonable form for computation can be. Define $W(S,t;K,r,\sigma^2)$ to be the Black-Scholes option pricing formula for the no-jump case given be equation 3.3. Define the random variable, X_n to have the same distribution as the product of n i.i.d. random variables, each identically distributed to the random variable Y, where it is understood that $X_0 = 1$. Define ϵ_n to be the expectation operator over the distribution of X_n. Then the solution to equation 30 when the current stock price is S can be written as

$$f(S,t) = \sum_{n=0}^{\infty} \frac{e^{-\lambda(T-t)}\lambda^n (T-t)^n}{n!} [\epsilon_n \{W(SX_n e^{-\lambda k(T-t)}, t; K, \sigma^2, r)\}]$$

While not a closed form solution, this equation is useful for computational purposes provided that the density function for Y is not too complicated.

6.5 Pricing Options for Processes Suggested by the Empirical Literature

The above option pricing theory and applications, hopefully, can be adapted to price options on price processes suggested by the empirical literature. The mixed normal distribution of Kon (1984), for instance, might be formulated as a stochastic process whose drift parameter and volatility are time dependent and shift among a finite number of values to be determined from real data. Once the process is specified, the Black-Scholes argument could be adapted to price the option. We conjecture that the correct option pricing formula in this case would be the expected value of the standard Black-Scholes formula conditioned on the distribution of the volatility parameter. This solution, if correct, would be computationally tractable because the values for σ^2 would be drawn from a discrete set.

Similarly, risk-neutral pricing methods might be used to compute the appropriate price of an option on an index or stock with returns that follow a t-distribution by discounting the expected value of the terminal option price. Options on stocks with returns governed by a truncated Levy flight might be treated in the same manner. Although we have omitted discussion here, Madan et al. (1998) treats the problem of pricing options when stocks follow a Variance Gamma process.

7 Conclusion

In this paper, we have given a careful proof of the Black-Scholes differential equation for pricing options, followed by a discussion of alternative candidates for the stochastic process of returns. Of the alternative processes discussed in the empirical literature, the discrete mixture of normal distributions model of Kon (1984) and the Variance Gamma (V.G.) model of Madan and Seneta (1990) seem quite promising. The option pricing problem for European options has been solved for the latter process, but to our knowledge no one has tackled the problem for the discrete mixture of normal distributions model. In addition, we are unaware of the existence of solutions to the option pricing problem for processes based on the scaled t-distribution for returns or the truncated Levy flight.

Besides adapting the method for pricing options to the alternative stochastic price processes mentioned above, this paper has several other empirical and theoretical extensions. First, no systematic study has been completed that compares the descriptive performance of the all the models for return distribution presented herein on a large sample of prominent US stocks and stock indices for recent data. A paper of this sort could help identify prominent candidates for replacing geomertic Brownian motion in a "bottom-up" fashion.

Second, a theoretically interesting extension of this work would be the examination of the effect of rate of convergence to Gaussian behavior (rather, to GBM) on the properties of option prices for suitable price processes. A variant of the truncated Levy flight might be adapted for this purpose. This project might yield insight, on a practical level, into the option pricing problem for stock options with long maturities, for which the convergence of the return distribution is a possibility.

Naturally, any price process suggested for replacing geometric Brownian motion must result in a formula that is comprehensible to option traders. Otherwise, they will not use it. Contingent upon developing option pricing formulas for some of the alternative processes mentioned in this paper, and in particular for the diffusion and jump processes that already allow closed form solutions, much work needs to be done to assess the success of these models in practical tasks such as delta hedging and minimizing transactions costs.

To do justice to the problem of alternative price processes, and in fleshing out the exposition in this paper, future additions to this particular work must include a more in-depth discussion of the theory of risk-neutral pricing, a treatment of the mathematics of diffusion processes and partial differential equations, and applications of the option pricing theory herein to some of the alternative candidates mentioned. Also, computational investigation of the properties of the option price distributions obtained from these processes would be very enlightening, as would further empirical investigation of the type mentioned above.

8 Acknowledgements

I would like to thank Prof. Arlie Petters and Brian Rider for their helpful advice on option pricing and preparing this paper.

This work is supported by NSF VIGRE grant number DMS-9983320.

References

[1] Almgren, R., 2002, Financial derivatives and partial differential equations, The American Mathematical Monthly 109, no. 1, 1-12.

[2] Bachelier, L., 1900, Theorie de la speculation, Annales Scientifiques de l'Ecole Normale Superieure 17, 21-86.

[3] Bakshi, G., C. Cao, and Z. Chen, 1997, Empirical performance of alternative option pricing models, The Journal of Finance 52, issue 5, 2003-2049.

[4] Black, F., and M. Scholes, 1973, The pricing of options and corporate liabilities, Journal of Political Economy 81, 637-659.

[5] Bookstaber, R.M. and J.B. McDonald, 1987, A general distribution for describing security price returns, Journal of Business 60 (July), 401-24.

[6] Breiman, L., 1968, Probability (Addison-Wesley, Reading, Massachusetts).

[7] Cox, J.C., and S. Ross, 1976, The valuation of options for alternative stochastic processes, Journal of Financial Economics 3, 145-166.

[8] Fama, E., 1965, The behaviour of stock market prices, Journal of Business 38 (January), 34-105.

[9] Feller, W., 1966, An introduction to probability theory and its applications, vols. I and II (Wiley, New York).

[10] Gopikrishnan, P., M. Meyer, L.A.N. Amaral, and H.E. Stanley, 1998, Inverse cubic law for the distribution of stock price variations, European Physics Journal B 3, 139-140.

[11] Kon, S.J., 1984, Models of stock returns–a comparison, Journal of Finance 39, issue 1, 147-165.

[12] Koponen, I., 1995, Analytic approach to the problem of convergence of truncated Levy flights towards the Gaussian stochastic process, Physical Review E 52, no. 1, 1197-1199.

[13] Madan, D.B., P.P. Carr, and E.C. Chang, 1998, The variance gamma process and option pricing, European Finance Review 2, 79-105.

[14] Madan, D.B., and E. Seneta, 1990, The variance gamma (V.G.) model for share market returns, Journal of Business 63, no. 4, 511-524.

[15] Mandelbrot, B.B., 1963, New methods in statistical economics, Journal of Political Economy 71, 421-40.

[16] Mantegna, R.N., and H.E. Stanley, 1994, Stochastic process with ultra-slow convergence to the Gaussian: the truncated Levy flight, Physical Review Letters 73, no. 22, 2946-2949.

[17] Merton, R.C., 1973, Theory of rational option pricing, Bell Journal of Economics and Management Science 4, 141-183.

[18] Merton, R.C., 1976, Option pricing when underlying stock returns are discontinuous, Journal of Financial Economics 3, 125-144.

[19] Officer, R.R., 1972, The distribution of stock returns, Journal of the American Statistical Association 67, issue 340, 807-812.

[20] Osborne, M., 1959, Brownian motion in the stock market, Operations Research 7, 145-173.

[21] Owen, J. and R. Rabinovitch, 1983, On the class of elliptical distributions and their applications to the theory of portfolio choice, Journal of Finance 38 (June), 745-52.

[22] Plerou, V., P. Gopikrishnan, L.A.N. Amaral, M. Meyer, and H.E. Stanley, 1999, Scaling of the distribution of price fluctuations of individual companies, Physical Review E 60, no. 6, 6519-?????

[23] Praetz, P.D., 1972, The distribution of share price changes, Journal of Business 45, issue 1, 49-55.

[24] Press, S.J., 1967, A compound events model for security prices, Journal of Business 40 (July), 317-335.