8. Option Valuation

8.1 Mathematical Background

Stochastic Processes: Basic Definitions[1]

The development of options pricing theory is intimately related to notions associated with ***stochastic processes***. The first important work on options pricing, Louis Bachelier's (1900) doctoral dissertation, also represents a significant early contribution to the theory of Brownian motion. Bachelier's work predates and anticipates Einstein's work on Brownian motion five years later. Unfortunately, Bachelier's thesis passed largely unnoticed and was only 'rediscovered' by Paul Samuelson around 1954, following the 'rediscovery' by Leonard Savage of a (1914) Bachelier publication on speculation and investment (Bernstein 1992). Bachelier entered the mainstream of financial economics in the mid-1960's when his thesis was included in Cootner's (1964) seminal book of readings on the random behavior of stock prices.[2]

The theory of stochastic processes, proper, has a much longer history. One possible starting point would be the work of Abraham de Moivre in the 1730's where he derived the normal distribution as the limit of the skew binomial. A more traditional starting point dates back to 1827 when the English botanist R. Brown observed that small particles, suspended in a liquid, exhibited "ceaseless irregular motions". This observation was subsequently applied to behavior of various other physical objects, such as smoke particles suspended in the air.[3] The important modern contributions in stochastic processes can be traced to N. Wiener in 1918.[4] His role is recognized in the use of the term ***Wiener process*** to signify the fundamental building block of the theory of stochastic differential equations. In recognition of the early work of Brown, the term ***Brownian motion*** is often used synonymously. Important later contributions were made by Kolmogorov and Feller. Excellent representations of the modern state of the theory can be found in Cox and Miller (1965), Gihman and Skorohod (1975, 1979), Arnold (1974), Karatzas and Shreve (1988), Karlin and Taylor (1981), and Rogers and Williams (1987).

Formally, a stochastic process can be defined:

> ***Definition***: Let $\{X(t)\}$ be a family of random variables indexed by the linear (index) set $\mathfrak{I}$, where $t \in \mathfrak{I}$. Then $\{X(t)\}$ is said to be a ***stochastic process.***

The terms stochastic (random) process and time series are often used interchangeably. Following Karlin and Taylor (1975, p.32), "a stochastic process may be considered as well defined once its state space, index parameter and family of joint distributions are prescribed." Similar approaches can be found in Goldberger (1964, p.142): "...a stochastic process is defined by a family of random variables $\{X(t)\}$....such that for all finite sets of choices of t ... a joint probability distribution is defined for the random variables $X(t)_1, X(t)_2,....,X(t)_n$"; and, Dhrymes (1974, p.383): "The probability characteristics of a stochastic process $\{X(t)\}$ are completely specified if we determine the joint density function of a finite number of members of the family of random variables comprising the process."

Heuristically, the theory of stochastic processes describes the behavior of random variables, the *X's*, over time, $t \in \mathfrak{I}$. A random variable is a function that maps from a prespecified domain, or sample space, to some portion of the real line, R^1. In certain financial applications, e.g., where X refers to an asset price, X takes values only on the positive, half line. In this case as well as when the X values are allowed to assume any value along the real line, it is conventional to assume that there is a zero probability of X being equal to plus or minus infinity. When t is fixed at a given point, $X(t)$ has the conventional interpretation of a random variable, with associated (one-dimensional) probability density function. Specification of the stochastic process for X requires further specification of the joint density functions that relate *X's* at different points in time: the joint densities provide a probabilistic specification of how X evolves over time. This potentially complicated mapping can involve various combinations of discrete or continuous observation on X and t. While in financial applications X is usually continuous, time can be either. The terms ***discrete stochastic process*** and ***continuous stochastic process*** are used to refer to the time intervals at which $X(t)$ is observed. Applications of stochastic processes have contributed significantly in numerous fields, e.g., engineering, physics, biology and statistics.

A number of important notions from stochastic processes have received considerable attention in financial economics. While martingales and random walks are possibly the most familiar,[5] most of the literature on options pricing is developed using diffusion processes. These concepts are all closely related. ***Diffusions*** are (strong) Markov processes, continuous in both X and t. In turn, diffusions are constructed from Wiener processes that are the continuous time representation of a discrete time random walk. Both types of processes obey the martingale property. The essential concept of a ***Markov*** process is most readily illustrated in discrete time. Using the conditional probability distribution $P[\cdot]$, then the Markov property can be stated:

$$P[\ X(t+1) = j \mid X(0)=a, X(1)=b, X(2)=c......., X(t)=s] = P[\ X(t+1) = j \mid X(t)=s]$$

In effect, the Markov process is "memory less", the past and future are statistically independent when the present is known.

While there are a number of subtle technical issues associated with the precise definition of Markov process, Markov families and diffusions (e.g., Wentzel 1981), for current purposes it is sufficient to define a diffusion process as a Markov process in which both time and the state variable X are continuous. When expressed in the form of a ***stochastic differential equation***, diffusions can be used to concisely specify the joint density functions for the stochastic process. To adequately motivate stochastic differential equations, it is necessary to develop the basic concept of a Wiener process as the limit of a discrete "normal" random walk, defined by a stochastic difference equation. The discrete ***random walk***, without drift, has the form:

$$X(t+1) = X(t) + Z(t+1) \quad \text{where } X(0) = 0, \text{ and } t \in \{1,2,3.....\}$$

where $Z(1), Z(2), Z(3)....$ form a stochastic process of independent random variables with the standard normal probability distribution: $Z(t) \sim N[0,1]$. In other words, over the unit time interval $(\Delta\, t = 1)\ Z(t)$ is a normal random variable with mean 0 and variance of one.

Types of Diffusions

Consider what happens to normal random walk as the time interval Δt shrinks. In this case, $Z(t + \Delta t)$ is no longer $N[0,1]$, but rather $N[0,\Delta t]$ where the random walk now has the form:

$$X(t+\Delta t) = X(t) + Z(t + \Delta t) = X(t) + Z(t)\,\sqrt{\Delta t}$$

Reexpressing in difference form and using:

$$\lim_{\Delta t \to 0} \Delta t = dt$$

Gives the stochastic differential equation for the standard Wiener process:

$$dX(t) = Z(t)\,\sqrt{dt} \quad \Rightarrow \quad dX(t) = dW(t)$$

$W(t)$ is usually referred to as a standard Wiener process with a unit variance parameter. The ***ensemble*** of sample paths for $X(t)$ conform to the evolution of a random variable that is standard normal on the unit time interval.

Geometrically, the behavior of the Wiener process can be illustrated in two dimensions by taking $X(t)$ on the vertical axis and time on the horizontal axis. Starting from $X(0)$, which is a predetermined point, the Wiener process specifies an infinite number of possible paths originating from $X(0)$. The pattern of these paths conform to $N[0,\Delta t]$. To see this, select any time $t_1 > 0$, take a 'slice' across the X paths and plot the distribution of the paths. The distribution of the paths will be a normal distribution, centered at $X(0)$, with variance t_1. Similarly, doing the same 'slicing' operation for another time $t_2 > t_1$ and evaluating the density associated with the ensemble of X time paths will again produce normal distribution, centered at $X(0)$, but with a larger dispersion. In this fashion, it is possible to make a connection with the distribution theory, familiar from traditional statistics, which is concerned with variables at a given point in time. The SDE technique is a method of describing the evolution over time of random variables that are function of the class of normal random variables.

The Wiener process can be immediately generalized to allow for non-zero drift and variance that differs from Δt. When a trend or "drift" term μ and standard deviation σ $(\neq 1)$ are admitted then the ***arithmetic Gaussian*** stochastic process (with constant coefficients μ and σ) is defined:

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

This constant coefficient process is also referred to as an arithmetic or absolute Brownian motion. This ensemble of time paths for this process are also normal but differ from the standard Wiener process by allowing for a different amount of variation around a constant trend. The basic Wiener process can be used to construct a wide range of stochastic differential equations, each of which is associated with a different specification for the joint density of the stochastic process. The construction of the Wiener process requires these densities to be functionally connected to the normal.

In general, the drift and standard deviation can be functions of both the state variable and time, such that $\alpha[x,t]$ is the drift and $\beta[x,t]$ is the volatility. Consider, a simple form of state dependence, where $\alpha[x,t] = \mu X$ and $\beta[x,t] = \sigma X$, with μ and σ being constants:

$$dX(t) = \mu \, X \, dt + \sigma \, X \, dW(t) \;\Rightarrow\; dX(t)/X = \mu \, dt + \sigma \, dW(t)$$

In this case, the instantaneous rate of change (dX/X) follows a Gaussian process. For this reason, the terms ***geometric Gaussian process*** or ***geometric Brownian motion*** are sometimes used to identify this process. It can be shown that for the geometric Gaussian process the paths of $X(t)$ correspond to a process that is log-normally distributed at each point in time. This is important for cases where $X(t)$ refers to prices that, like log-normal variables, cannot be negative (for $X(0) > 0$).[6]

The geometric Brownian motion has an important position in option pricing theory. Black and Scholes (1973) use the geometric Gaussian process to describe the behavior of the stock price in deriving their option pricing formula. The process is often encountered in a variety of option pricing situations because geometric Brownian motion usually leads to a relatively simple closed form solution. While the empirical validity of this assumption can be questioned, the process is sufficiently close to real world processes that in many cases a reasonable approximation is provided.[7] In addition, recent work by Nelson (1990) has shown that allowing for $\mu = \mu[X]$ and $\sigma = \sigma[X]$ produces the result that the geometric Gaussian process is the continuous limit of the popular GARCH discrete stochastic process.

For various reasons, it is useful to know the mean and variance of an assumed diffusion process. For example, a key step in the risk neutral valuation of the European option requires this result. For geometric Brownian motion:

$$E[X(t)] = X(0) \, e^{\alpha t} = X(0) \, e^{(\mu + \frac{\sigma^2}{2}) t} \qquad var[X(t)] = X(0)^2 \, e^{2 \alpha t} \, [e^{\sigma^2 t} - 1]$$

While not difficult, the derivation of these results is left to Appendix III.

Other important processes that are exploited in the pricing of options include variations of the ***regular Ornstein-Uhlenbeck*** (OU) process:[8]

$$dX(t) = \alpha \, X \, dt + \sigma \, dW(t)$$

In financial applications, the OU process is often presented as a ***mean reverting*** process:

$$dX(t) = \alpha[\mu - X]\, dt + \sigma\, dW(t)$$

where α can be interpreted as the speed of adjustment of $X(t)$ to the steady state mean μ. Different variations of the OU process have appeared. For example, Brennan and Schwartz (1979) and many others use the process:

$$dX(t) = \alpha[\mu - X]\, dt + \sigma\, X\, dW(t)$$

In this case, dX/X reduces to a mean-reverting OU.

Unlike the geometric and arithmetic Brownian motions, the $X(t)$ distribution associated with the OU process changes over time. At any time t, $X(t)$ is governed by a non-steady state distribution that is normal but with ***conditional*** means and variances that are time dependent, e.g., Cox and Miller (p.225-27). For the ***regular OU:***

$$E[X(t)] = X(0)\, e^{-\alpha t} \qquad var[X(t)] = \frac{\sigma^2\,(1 - e^{-\alpha t})}{2\,\alpha}$$

Asymptotically, as $t \to \infty$ then the OU converges to a normal steady state distribution with mean zero and variance $\sigma^2/2\alpha$. This damping behavior of the OU is different from the geometric Brownian motion that increases indefinitely as $t \to \infty$. Similarly, the ***mean-reverting OU*** has conditional parameters:

$$E[X(t)] = \mu + (X(0) - \mu)\, e^{-\alpha t} \qquad var[X(t)] = \frac{\sigma^2}{2\alpha}\, \{1 - e^{-2\alpha t}\}$$

As $t \to \infty$ then the mean-reverting OU converges to a steady state normal distribution with mean μ and variance $\sigma^2/2\alpha$. Cox, Ingersoll and Ross (1985) and others use a diffusion process, the square root process, in which volatility depends on the square root of X. In mean-reverting form, the SDE for this process is represented:

$$dX(t) = \alpha[\mu - X]\, dt + \sigma \sqrt{X}\, dW(t)$$

This process can be shown to be the continuous limit of a ***non-central*** chi-squared distribution. In turn, the square root process is a special case of the more general ***constant elasticity of variance*** (CEV) process:

$$dX(t) = \alpha X\, dt + \sigma X^{\beta/2}\, dW(t)$$

where $\beta \in [0,2)$. Unlike the other diffusions examined to this point, the CEV model describes a class of processes defined over the range of β. For $\beta = 1$, the CEV process is equivalent to a square root process. For $\beta = 2$, the process is lognormal.

The ***Brownian bridge*** process is a development on the Brownian motion process that incorporates a drift term which forces the process to a fixed endpoint. For a Brownian bridge process that starts at zero and ends at zero, the unit variance process takes the form:

$$dX(t) = -X(t)/(T - t)\, dt + dW(t)$$

The Brownian bridge process has been used in a number of papers, e.g., Ball and Torous (1983), Chiang and Okunev (1993), to model the bond price process. For this purpose, the Brownian bridge is attractive because the bond price converges to par value at the term to maturity endpoint. Cheng (1991) explores problems that assuming a Brownian bridge has for arbitrage free pricing.

Unfortunately, the number of diffusion processes for which readily interpretable conditional probability densities can be derived is limited. While the density function for the general CEV process has been derived, the formula is complicated. In addition to the processes already described and immediate extensions, such as the Brownian bridge, the only other class of processes that have manageable density functions are Bessel processes which, in squared form, have the representation: $dX(t) = \alpha\ dt + 2\sqrt{X}\ dW(t)$. This form of the Bessel process is typically used, e.g., Geman and Yor (1993), because the associated densities are stable under convolutions, a useful property not shared by geometric Brownian motion. While not always the best empirical fit to the distribution for financial random variables, the analytical advantages of diffusion processes have played important roles in the application of stochastic processes to problems in finance and economics.[9]

With relatively little analytical complication, the general SDE model can be extended to include ***jump processes***. One of the properties of the diffusion is the continuity of its sample paths. This property can be generalized to permit certain types of jumps in the state variable to take place, usually modeled as a Poisson event process. In other words, the sample paths of the diffusion are continuous except at a countable number of discontinuity points, where the jumps are generated by a Poisson process. In this case, the SDE can be written:

$$dX(t) = \alpha[x,t]\ dt + \sigma[x,t]\ dW(t) + v[x,t]\ dQ(t)$$

where $v[x,t]$ obeys the same technical conditions imposed on α and σ, and $dQ(t)$ is a Poisson process assumed to be distributed independently of $W(t)$.[10] For this type of SDE, there is a generalized form of Ito's lemma, e.g., Malliaris and Brock (1982, p.122). The usefulness of this result in applications is that it, theoretically, permits the valuation of jumps in the underlying process, e.g., Merton (1976), Cox and Ross (1976).

Another complication that can be introduced to the SDE is to impose ***absorbing*** or ***reflecting barriers*** on the process. An absorbing barrier occurs when the process vanishes at a specific point. A natural example of this occurs for price processes that vanish when the price process reaches zero. The absorbing barrier is imposed by requiring the transition probability density to equal zero when the state variable equals the absorbing value. The resulting solution for the absorbed transition density typically has a solution that is the difference of the unrestricted density and the density associated with the paths that are absorbed, e.g., Karlin and Taylor (v.1, p.355). Situations where a reflecting barrier is imposed on the SDE have less immediate natural application in Finance. The imposition of a reflecting barrier involves restrictions on the partial derivative of the transition density with respect to the state variable. The restricted transition density will be the difference between the unrestricted density and the density associated with the transients induced by the reflecting barrier, e.g., Cox and Miller (1965, p.224).

The Chapman-Kolmogorov Equations

Associated with a general SDE of the form $dX(t) = \alpha[X,t]\ dt + \sigma[X,t]\ dW(t)$, together with an initial condition $X(0) = c$, is an integral solution for $X(t)$:

$$X(t) = X(0) + \int_0^t \alpha[s]\ ds + \int_0^t \sigma[s]\ dW(s)$$

Considerable analysis has been done on identifying the conditions under which $X(t)$ can be derived from the SDE, e.g., Malliaris and Brock (1982, Chap. 2). The existence and uniqueness of such a solution depends on a number of technical conditions being satisfied, primarily associated with the behavior of α and σ.[11] Heuristically, the most commonly used form of these conditions, the Lipschutz and growth conditions, can be interpreted as restrictions that the coefficients cannot grow faster than the random variables.

Given a specific solution, $X(t)$, for the SDE, a number of important analytical results apply. One essential feature of a diffusion process is that the transition probability density, the probability law that determines how $X(t)$ evolves through time, will be Markov. If $X(t)$ is a solution to the SDE then, subject to the satisfaction of a number of technical conditions, the transition probability density $P[X_0,t_0;X,t]$ can be obtained from α and σ using the Kolmogorov forward and backward equations that are PDE's which the transition probability density obeys. For the ***arithmetic Gaussian*** process, that features constant coefficients, the forward or Fokker-Planck equation is:

$$\frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2}P[X_0,t_0;\ X,t]\ -\ \alpha\frac{\partial}{\partial x}P[X_0,t_0;\ X,t]\ =\ \frac{\partial}{\partial t}P[X_0,t_0;\ X,t]$$

In financial applications, this PDE equation is usually subject to the initial condition $P[X_0,t_0;\ X,t] = \delta(X - X_0)$.

The forward equation is so-named because the solution process starts from the initial values X_0 and t_0 and solves for the density function of the future values of $X(t)$. This is the usual method of solving valuation problems in finance. The forward equation for transition densities that have non-constant coefficients is:

$$\frac{1}{2}\frac{\partial^2}{\partial x^2}\{\sigma^2[X,t]\ P[X_0,t_0;\ X,t]\}\ -\ \frac{\partial}{\partial x}\{\alpha[X,t]\ P[X_0,t_0;\ X,t]\}\ =\ \frac{\partial}{\partial t}P[X_0,t_0;\ X,t]$$

Related to the forward equation is the backward equation, that is similar in appearance to the forward equation, and is used to derive the density function of X_0 starting from the density associated with $X(t)$. A demonstration of how to apply the forward equation to solve for specific functional form of a transition probability density for the OU and other processes in given in Cox and Miller (1965) and other sources.

7.2 Ito's Lemma

Univariate Ito's Lemma

Calculus is an important mathematical technique with numerous useful applications. As conventionally presented, calculus is applied to functions that are deterministic. In effect, the familiar rules associated with dy/dx, such as $y = x^2 \rightarrow dy/dx = 2x$, only apply when x is known. When x is a random variable the usual rules of calculus no longer apply. The importance of Ito's lemma is that it specifies the procedures for applying calculus to functions which contain random variables. More precisely, Ito's lemma provides a method for evaluating the total derivative of a function of a stochastic variable that follows a Markov diffusion process. As discussed in 7.1, the diffusion class includes a wide range of stochastic processes. Given this, the ***univariate*** form of Ito's Lemma can be stated:[12]

> ***Ito's Lemma***: Let $u(x,t)$ be a continuous random function mapping from $R^1 \times [0,T] \rightarrow R_1$ with continuous partial derivatives u_t, u_x, and u_{xx}. If $x(t)$ is a random process with a stochastic differential equation obeying a diffusion of the form:
>
> $$dx(t) = a(t)\,dt + v(t)\,dW(t)$$
>
> where $W(t)$ is a standard Wiener process and $a(t)$ and $v(t)$ are the drift and volatility of the diffusion, then the function $y(t) = u(x(t),t)$ also has a differential on $[0,T]$ given by:
>
> $$dy(t) = \{u_t + u_x\,a(t) + 1/2\,u_{xx}\,v(t)^2\}\,dt + u_x\,v(t)\,dW(t)$$

This form of Ito's lemma generalizes in a natural fashion to the case where x is multidimensional, e.g., Malliaris and Brock (p. 85-6).[13]

While Ito's lemma has had numerous applications in financial economics, perhaps the most well-known is the Black-Scholes application to call option valuation. In this case, the non-dividend paying stock price is assumed to follow a log-normal diffusion:

$$dS = \alpha\,S\,dt + \sigma\,S\,dW$$

In this case, $\boldsymbol{a(t) = \alpha S}$ **and** $\boldsymbol{v(t) = \sigma S}$ where α and σ are constants. The functional relationship between the call option price (C) and the stock takes the form: $C = C[S,t;\ X]$. Application of Ito's lemma gives:

$$dC = \{C_t + C_S\, \alpha\, S + 1/2\, C_{SS}\, \sigma^2 S^2\}\, dt + \{C_S\, \sigma\, S\}\, dW$$

This solution plays a central role in the derivation of the Black-Scholes price of a European call. In particular, Black-Scholes are able to use the ***riskless hedge portfolio*** construction to provide an additional condition that permits elimination of the *dW* term. In this fashion, the call option pricing problem is transformed from an SDE problem that is not solvable in closed form with standard techniques, into a deterministic partial differential equations (PDE) problem that can be solved. However, complications associated with solving PDE's are such that closed form solutions are not always possible. This has at least two implications. Firstly, there is an emphasis on problem specifications that can be solved in closed form, e.g., Black-Scholes, even though such solutions may not be fully realistic. And, secondly, there is the need to apply numerical simulation to 'solve' problems in which a precise specification of the problem is required.

Ito's lemma can also be used to transform one type of SDE into another. One useful example involves the function $G(S) = ln[S]$ when the SDE for *S* is geometric Brownian motion. Application of Ito's lemma provides the SDE for *G(S)*, specified using the drift and volatility parameters for S. More precisely, $G_t = 0$, $G_S = 1/S$ and $G_{SS} = -(1/S)^2$. Using Ito's lemma it follows that:

$$dG = \{0 + \alpha - 1/2\, \sigma^2\}\, dt + \sigma\, dW = (\alpha - 1/2\, \sigma^2\}dt + \sigma\, dW$$

The result that the log of a log-normally distributed random variable follows an arithmetic Brownian process is not surprising, but the specification of the drift is not obvious.

A straight forward extension to the case of functions of two variables (and time) is provided by Fischer (1975) which examines the stochastic behavior of the real bond price $q = B/P$ where B is the nominal price of a riskless bond and *P* is the price level. In this case, the rate of change in the price level *(dP/P)* and the return on a zero-coupon continuously compounded nominal bond *(dB/B)* can be specified:

$$dP/P = \pi\, dt + \sigma\, dW \qquad dB/B = R\, dt$$

Observing that when $q = u(B,P) = B/P$:

$$u_t = 0 \quad u_{BP} = -\{1/P\}^2 \quad u_{BB} = 0 \quad u_B = 1/P \quad u_P = -\{B/P^2\} \quad u_{PP} = 2\{B/P^3\}$$

It is now possible to apply the multivariate form of Ito's lemma to get:[14]

$$dq = \{ \frac{1}{P} RB + \frac{-B}{P^2}\pi P + \frac{B}{P^3}P^2\sigma^2 \}\, dt + \frac{-B}{P^2}\sigma P\, dW$$

Factoring out *B/P* gives the desired result:

$$dq/q = \{R - \pi + \sigma^2\}\, dt - \sigma\, dW$$

More developed applications of the multivariate form of Ito's lemma can be found in various sources, e.g., Gibson and Schwartz (1990) and Schwartz (1982).

Multivariate Ito's Lemma

It is useful to proceed by example to illustrate the multidimensional form of Ito's Lemma. Briys and Solnik (1992) provide an interesting application to the case where *x* is multidimensional, involving two random variables. This requires a more involved form of Ito's lemma. The Briys and Solnik example involves the domestic currency value of a foreign asset, $V^* = VS$, where *V* is the random foreign currency value of the foreign asset and *S* is the random spot exchange rate, producing V^* which is the random domestic currency value of the foreign asset. In this case, processes are given for dV and dS with dV^* to be calculated using Ito's lemma.

To evaluate Ito's lemma for this case requires application of the associated multivariate total derivative. The resulting solution for the two random variable case, $y(t) = u[t, \{x(t)\}] = u[t, \{x_1(t), x_2(t)\}]$, takes the form:

$$dy = u_t\, dt + u_x\, dx + \frac{1}{2}\, tr[\sum\ u_{xx}]\ dt$$

where Σ is the variance-covariance matrix of the state variables and $tr[\cdot]$ is the trace operator. Because there are n state variables in the general form, u_x is a $n \times 1$ column vector containing the first partial derivatives, dx is a row vector containing the diffusion processes and u_{xx} is a symmetric $n \times n$ matrix of the second partial derivatives.

In the Briys and Solnik example, $dy = dV^*$ and V and S are assumed to follow the log-normal diffusions:

$$\frac{dV}{V} = \mu_V\, dt + \sigma_V\, dW_V \qquad\qquad \frac{dS}{S} = \mu_S\, dt + \sigma_S\, dW_S$$

The covariance between dW_S and dW_V per unit time, $cov(dW_S, dW_V) = \rho_{VS}\, dt$. It follows that $\sigma_{VS} = \rho_{VS}\, \sigma_V\, \sigma_S$. Recognizing that $y = u(x,t)$ in this case is $V^* = VS$, it follows that:

$$u_1 = \frac{\partial V*}{\partial V} = S \qquad u_2 = \frac{\partial V*}{\partial S} = V \qquad u_t = \frac{\partial V*}{\partial t} = 0$$

$$u_{11} = \frac{\partial^2 V*}{\partial V^2} = 0 = \frac{\partial^2 V*}{\partial S^2} = u_{22} \qquad \frac{\partial^2 V*}{\partial V \partial S} = 1 = \frac{\partial^2 V*}{\partial S \partial V} = u_{1,2}$$

Because there are now two state variables, V and S, the variance-covariance matrix Σ *is* 2×2 with variances on the diagonal and covariances on the off-diagonal. Remembering from the univariate log-normal example that $v(t)^2 = \sigma^2 S^2$, evaluating $1/2\ tr[\cdot]$ for this case gives:

$$\frac{1}{2}\, tr\{\begin{bmatrix} \sigma_{11} & \sigma_{1,2} \\ \sigma_{2,1} & \sigma_{22} \end{bmatrix} \begin{bmatrix} u_{11} & u_{1,2} \\ u_{2,1} & u_{22} \end{bmatrix}\} = \frac{1}{2}\{\sigma_{11}u_{11} + \sigma_{1,2}u_{2,1} + \sigma_{2,1}u_{1,2} + \sigma_{22}u_{22}\} = \sigma_{1,2} = \sigma_{VS}\ VS$$

Using the result that $1/2\ tr[\cdot] = \sigma_{VS}\, VS$. The solution to the total derivative is found to be:

$$dV* = 0 + S(\mu_V V\, dt + \sigma_V V\, dW_V) + V(\mu_S S\, dt + \sigma_S S\, dW_S) + \sigma_{VS} VS\, dt$$

$$\frac{dV*}{V*} = (\mu_V + \mu_S + \sigma_{VS})\, dt + \sigma_V\, dW_V + \sigma_S\, dW_S$$

Given dV and dS, this is dV^*, the SDE for the domestic currency value of the foreign asset.

8.2 Deriving the Black-Scholes Option Pricing Formula[15]

Deriving the Formula

The Black-Scholes Assumptions

The Black-Scholes assumptions are:

a) Non-dividend paying stock.
b) European option.
c) The instantaneously riskless continuous interest rate r is constant over time (with a flat term structure).
d) The model has only ***one*** source of randomness, the single state variable, the price of the stock that follows a log-normal diffusion process. This log normal process is defined only over $S \in [0,\infty]$.
e) No transactions costs or taxes.
f) No penalties on short selling.
g) Riskless lending and borrowing at r.
h) Continuous trading.

Assumptions e)-g) are the conventional perfect markets assumptions.

The seminal paper by Black and Scholes (1973) directly addresses the problem of providing a practical, closed form expression for the price of a call option on a common stock prior to the maturity date. Extensions to options on futures and spot commodities follow with appropriate adjustments. Recognizing that a considerable amount of research effort preceded the development of Black-Scholes, e.g., Cootner (1964), what is most significant about this contribution is that the formula depends only on observable inputs: the stock price, the exercise price, time to expiration, the interest rate and the variance of the stock price. With suitable choice of estimator for the variance, the accuracy of the formula has proved to be robust across a wide range of option pricing situations. Extensions of this option pricing approach can be used to value: the debt and common stock of the firm; bond features such as callability, convertibility and retractability; and, tax and investment policy. As initially derived, the formula applies to a "perfect markets" world for a European call on a non-dividend paying stock. The techniques used to derive the formula are consistent with the material on stochastic processes covered previously. With appropriate transform of variables, Black and Scholes were able to show that the call option pricing problem reduces to solving a special case of the heat equation from mathematical physics.

The Black-Scholes argument proceeds by implementing the essential concept of a ***hedge portfolio***. Recognizing that there is both a long and a short hedge portfolio, the long portfolio involves combining a long stock position with α written call options on the stock. The hedge is created by selling just enough calls to offset changes in the value of the stock position. This portfolio will involve a net investment of funds because the premium income received from writing the calls will not be sufficient to purchase the stock position. Similarly, the short hedge portfolio involves shorting the stock, buying call options to hedge the position against upward changes in the stock price, and investing the balance of the funds in a riskless asset. Observe that, for both the long and short hedge portfolio, as the stock price changes the option hedge has to be continuously adjusted to maintain the hedge position.

For ease of exposition, consider the long hedge portfolio. In order to determine the number of call options to sell, let $V = S - \beta C$ be the value of the hedge portfolio. From the hedge portfolio construction:

$$\frac{\partial V}{\partial S} = 1 - \beta \frac{\partial C}{\partial S} = 0 \qquad \Rightarrow \quad \beta = \frac{1}{\dfrac{\partial C}{\partial S}} \qquad \Rightarrow \quad \frac{\partial C}{\partial S} = \frac{1}{\beta}$$

Given this specification for β, the change in the value of riskless hedge portfolio will earn the riskless rate of interest. (This follows from the assumption that interest rates are riskless). In terms of arbitrage portfolios, this condition is necessary in order to prevent the execution of arbitrage trades by either borrowing at the riskless rate and buying the hedge portfolio or selling the hedge and investing the funds at the riskless rate. It follows that there are two dV conditions that have to be satisfied:

$$\{S - \frac{C}{C_S}\} \; r \; dt = dV \qquad \Rightarrow \quad dV = dS - \frac{dC}{C_S}$$

At this point, Ito's lemma can be applied to solve for dC:

$$dC = \{C_t + C_S \, \alpha S + \frac{1}{2} C_{SS} \, \sigma^2 \, S^2\} \; dt + C_S \, \sigma \, S \; dW$$

The solution for dV can be derived as:

$$dV = dS - \frac{dC}{C_S}$$

$$= \alpha S \; dt + \sigma S \; dW - \frac{1}{C_S}\{[C_t + C_S \, \alpha S + \frac{1}{2} C_{SS}\sigma^2 \, S^2] \; dt + C_S \sigma S \; dW\}$$

$$= -\frac{1}{C_S}\{C_t + \frac{1}{2} C_{SS}\sigma^2 S^2\}dt = (S - \frac{C}{C_S})r \; dt$$

Dividing through by $-\{dt/C_S\}$ gives the ***fundamental partial differential equation*** for a European call on a nondividend paying stock:

$$C_t = rC - rS \; C_S - \frac{1}{2} \sigma^2 \, S^2 \; C_{SS}$$

With appropriate transformation of variables, this equation reduces to the heat equation.

The Black-Scholes option pricing formula is derived by solving the fundamental PDE, subject to appropriate boundary and terminal conditions: $C[S,0,X] = Max[0, S(T) - X]; \; C[0,\tau,X] = 0$. If Laplace transforms are used in the solution procedure, the condition $C_S[\infty,\tau,X] = 1$ is also used. As it turns out, solving the PDE problem involves considerable analytical manipulation. The ultimate solution is the Black-Scholes formula:[16]

$$\boldsymbol{C[S,t^*;X,r,\sigma] = C = S \; N[d_1] - X \; e^{-rt^*} \; N[d_2]}$$

where $N[\cdot]$ is the ***cumulative*** standard normal distribution evaluated at the appropriate argument and:[17]

$$d_1 = \frac{\ln[\frac{S}{X}] + (r + \frac{1}{2}\sigma^2)t^*}{\sigma\sqrt{t^*}}$$

$$d_2 = d_1 - \sigma\sqrt{t^*} = \frac{\ln[\frac{S}{X}] + (r - \frac{1}{2}\sigma^2)t^*}{\sigma\sqrt{t^*}}$$

The formula for pricing puts follows from put-call parity:

$$P = C + X\,e^{-rt^*} - S = S\,N[d_1] - X\,e^{-rt^*}\,N[d_2] + X\,e^{-rt^*} - S$$

$$= S\,\{N[d_1] - 1\} - X\,e^{-rt^*}\,\{N[d_2] - 1\} = X\,e^{-rt^*}\,N[-d_2] - S\,N[-d_1]$$

This follows from observing that $N[d] - 1 = N[-d]$ (see Appendix 3).

Manual evaluation of the Black-Scholes formula requires familiarity with the calculation of specific values for $N(d)$, the cumulative normal distribution function evaluated at d (see Table 8.1). Though it is easiest to work directly with $N(d)$, it is also possible to work with $n(d)$, the normal probability density function and to construct $N[d]$ values from n[d] tables. This is possible because the value of the cumulative normal distribution function evaluated at x, $N[x]$, is the area under the normal density function between $-\infty$ to x.

To evaluate the Black-Scholes formula using the standard normal probability density table, refer to Tables 8.1 and 8.2 that have both $N[d]$ and $n[d]$. The potentially simple exercise of producing the required cumulative $N[d]$ values is complicated by the form in which the $n[d]$ table is presented. To illustrate the use of the table, observe that when the calculated value d in $N[d]$ is greater than zero, the value in the $n[d]$ table is added to $.5 = N[0]$. Similarly, when $d < 0$, then the value in the $n[d]$ table is subtracted from .5. In the example provided, d_1 was approximately -.25. To identify the appropriate $N[\cdot]$ value, look down the vertical axis in the $n[d]$ table to find the first digit, in this case .2, and then read across to get the second digit, in this case .05. Examining the table the value in that cell is .0987. Subtracting this value from .5, because $d_1 < 0$, gives .4013, the value used in the example. Evaluating d_2 *at* $-.50$ gives a Table value of .1915 and $N[d_2] = .3085$. These are the same values as those obtained by working directly from $N[d]$. This method of calculating d using the relevant standard normal densities can be used because the distribution is symmetric, making the values for one half of the distribution redundant. Presenting the table using only half of the distribution information as in the Table allows a larger number of 'useful' cells to be presented.

Interpreting the Formula

The first step in interpreting the Black-Scholes formula involves interpreting $N[d]$, particularly as S and X change. Consider an at-the-money option, $S=X$, where t^* is small. In this case, ln $[S/X] = ln[1] = 0$ and, due to small t^*, the remaining values will be small and $N[d] \rightarrow N[0] = .5$. Similarly, when the option is deep in-the-money, $S >> X$ *and* $N[d] \rightarrow N[+\infty] = 1$, which implies that, for an European call option on a non-dividend paying security, $C \rightarrow S - X\,exp[-rt^*] > S$ - X. Observing that $S - X$ is the early exercise value, it follows that European call options on non-dividend paying securities will, if the Black-Scholes formula is correct, not be exercised early. For puts, deep in-the-money requires $X >> S$ where $N[-d] \rightarrow N[+\infty]$, which implies for a put option that: $P \rightarrow X$ exp$[-rt^*]$ - $S < X - S$. It follows that European puts on non-dividend paying securities can be rationally exercised early. This asymmetry is an artifact of the upper bound on the value of the put position resulting from the restriction that $S \geq 0$. If the spot price falls far enough, the expected loss of interest due to early exercise will exceed the expected possible gains due to further price declines.

It has been demonstrated that as an option goes deeper into the money, the ln $[S/X]$ term in d_1 and d_2 tends to dominate the value of d. The same is true as the options goes deeper out of the money. For the call option, deep out-of-the-money requires $X >> S$. This produces: $N[d] \rightarrow N[-\infty]$ because $ln[0] = -\infty$. Hence, as the spot prices goes to zero then: $C \rightarrow 0$. A similar result holds for puts, with the proviso that deep out-of-the-money requires $S >> X$ for puts. In this case $N[-d] = N[-(+\infty)] = 0$. Similar intuition can be used to evaluate the formula as $t^* \rightarrow 0$. In this case, only the term involving ln $[S/X]$ matters, due to the t^* cancellation of $\sqrt{t^*}$ leaving a $\sqrt{t^*}$ term to multiply $(r + .5\ \sigma^2)$. If the option is in-the-money on the expiration date than $C = S - X$, whereas if the option is out-of-the-money then $C = 0$. This is a verification that the formula satisfies the boundary condition that: $C[X,t{=}T] = max[S - X, 0]$.

Other basic interpretations of the formula are associated with the relevant partial derivatives. Analytically, the partial derivatives:

$$\frac{\partial C}{\partial t} \qquad \frac{\partial C}{\partial S} \qquad \frac{\partial^2 C}{\partial S^2}$$

of the fundamental PDE, together with other partial derivatives of the Black-Scholes formula, are demonstrated in Sec. 8.2 to play a central role in applications to the management of portfolios. As it turns out, for portfolios containing derivative securities, one or more of the partial derivatives become choice variables. To illustrate possible applications, consider the riskless hedge portfolio. In this case, the delta of the portfolio is constructed to have $\partial V/\partial S = 0$.

Another simple example is a portfolio containing only a call option. By setting, say, $\partial^2 C/\partial S^2 = 0$ by choosing to write only options that are deep in-the money, the fundamental PDE provides restrictions on the remaining two partial derivatives of the form:

$$\frac{\partial C}{\partial t} + rS\,\frac{\partial C}{\partial S} = rC$$

Recognizing that $\partial C/\partial S = 1$ when the option is deep in-the-money provides a solution for $\partial C/\partial t$ in terms of the interest expense of the funds invested in the position. Chapter 10 provides numerous more developed extensions of this approach to the construction of portfolios containing options.

Application of the Black-Scholes Formula

Assume the following: $S(t)$ = \$36, X = \$40, τ = 3 months $\Rightarrow$ t^* = .25, r = .05, σ = .5. Both the interest rate and standard deviation are expressed in annualized form. For sigma, this requires estimating the standard deviation over the relevant sampling frequency and then annualizing as appropriate. Given this:

$$d_1 = \{\ln[36/40] + (.05 + .5(.5)^2)\ .25\} \,/\, \{.5\ (\sqrt{.25})\} \cong -.25$$

$$d_2 = \{\ln[36/40] + (.05 - .5(.5)^2)\ .25\} \,/\, \{.5\ (\sqrt{.25})\} \cong -.50$$

Evaluating the $N[\cdot]$ values: $N[-.25]$ = .4013 and $N[-.50]$ = .3085 and exp{(.05)(.25)} = .9877, it is possible to solve for the Black-Scholes call option price:

$$C = S\ [.4013] - X\ [.9877]\ [.3085] = 14.4468 - 12.1882 = \$2.26$$

Table 8.1 **Values of the Cumulative Normal Distribution Function**

Values of $N(d)$ for Selected Values of d

d	$N(d)$	d	$N(d)$	d	$N(d)$
		−1.00	.1587	1.00	.8413
−2.95	.0016	−.95	.1711	1.05	.8531
−2.90	.0019	−.90	.1841	1.10	.8643
−2.85	.0022	−.85	.1977	1.15	.8749
−2.80	.0026	−.80	.2119	1.20	.8849
−2.75	.0030	−.75	.2266	1.25	.8944
−2.70	.0035	−.70	.2420	1.30	.9032
−2.65	.0040	−.65	.2578	1.35	.9115
−2.60	.0047	−.60	.2743	1.40	.9192
−2.55	.0054	−.55	.2912	1.45	.9265
−2.50	.0062	−.50	.3085	1.50	.9332
−2.45	.0071	−.45	.3264	1.55	.9394
−2.40	.0082	−.40	.3446	1.60	.9452
−2.35	.0094	−.35	.3632	1.65	.9505
−2.30	.0107	−.30	.3821	1.70	.9554
−2.25	.0122	−.25	.4013	1.75	.9599
−2.20	.0139	−.20	.4207	1.80	.9641
−2.15	.0158	−.15	.4404	1.85	.9678
−2.10	.0179	−.10	.4602	1.90	.9713
−2.05	.0202	−.05	.4801	1.95	.9744
−2.00	.0228	.00	.5000	2.00	.9773
−1.95	.0256	.05	.5199	2.05	.9798
−1.90	.0287	.10	.5398	2.10	.9821
−1.85	.0322	.15	.5596	2.15	.9842
−1.80	.0359	.20	.5793	2.20	.9861
−1.75	.0401	.25	.5987	2.25	.9878
−1.70	.0446	.30	.6179	2.30	.9893
−1.65	.0495	.35	.6368	2.35	.9906
−1.60	.0548	.40	.6554	2.40	.9918
−1.55	.0606	.45	.6736	2.45	.9929
−1.50	.0668	.50	.6915	2.50	.9938
−1.45	.0735	.55	.7088	2.55	.9946
−1.40	.0808	.60	.7257	2.60	.9953
−1.35	.0885	.65	.7422	2.65	.9960
−1.30	.0968	.70	.7580	2.70	.9965
−1.25	.1057	.75	.7734	2.75	.9970
−1.20	.1151	.80	.7881	2.80	.9974
−1.15	.1251	.85	.8023	2.85	.9978
−1.10	.1357	.90	.8159	2.90	.9981
−1.05	.1469	.95	.8289	2.95	.9984

Table 8.2 **Value of the Normal Density Function**

THE STANDARD NORMAL DISTRIBUTION

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Note: For $x > 0$, add .5 to the value in the Table. For $x < 0$, subtract the value in the Table from .5

Implied Volatility

Malz (2000) on OTC Foreign Currency Option Trading

In foreign exchange markets, options are traded among dealers using the delta as a metric for exercise price and the Black-Scholes volatility as a metric for price. The volatility smile can then be represented as a schedule of implied volatilities of options on the same underlying and with the same maturity but different deltas. This convention is unambiguous, since a unique exercise price corresponds to each call or put delta and a unique option premium in currency units corresponds to each implied volatility. The most liquid option markets are for at-the-money forward and for 25-delta calls and puts. There are also somewhat less actively traded markets in 10-delta calls and puts.

Practical application of the Black-Scholes formula requires some method of estimating the variance, σ^2, of the continuously compounded rate of return on the stock *(dS/S)*.[18] Unlike the other values required to calculate the Black-Scholes option price, (S,X,r,τ), the value for volatility is not directly observed. Some method of estimating σ is required. Starting with the first empirical studies on Black-Scholes, it has been recognized that using the past history of returns to estimate volatility can produce undesirable results, e.g., inconsistency between Black-Scholes and observed option prices, e.g., Latane and Rendleman (1976), Macbeth and Merville (1979).[19] Lauterbach and Schultz (1990) reports: "...our findings indicate that the constant equity variance assumption is the most serious deficiency in the Black-Scholes model." Over time it has been recognized that, given an observed call option price, Black-Scholes can be "inverted" and used to solve for an estimate of volatility. In other words, given S,X,r,τ **and C** it is possible to use Black-Scholes to solve for σ. For this purposes, computer programs that solve for Black-Scholes prices also provide a routine for solving the *implied standard deviation* or implied volatility (IV).[20]

Not surprisingly, the behavior of implied volatility has been intensely studied from a number of different angles. From the beginning, study of implied volatility has revealed that Black-Scholes is not a formula for pricing options but, rather, because the traded call price is observed and the only unobserved variable in the Black-Scholes model is the volatility, Black-Scholes is more appropriately viewed as a formula for determining an estimate of the underlying spot price *(S)* volatility.[21] Various properties of the implied volatility have been examined in numerous studies. One property of interest is the ability of the implied volatility to predict future volatility, typically using the most advanced econometric volatility predictor, e.g., GARCH, as a benchmark. This type of comparison can be found in Stein (1989), Day and Lewis (1992), Canina and Figlewski (1993), Lamoureux and Lastrepes (1993), Heynen, Kemna and Vorst (1994), Takezawa (1995) and Xu and Taylor (1994, 1995), Sabatini and Linton (1998), Dumitru (1999).

Volatility forecasts have a number of potential uses, including being used as input for market markers determining appropriate bid/offer quotes for option trading or as inputs to Value at Risk calculations. For a number of reasons, it is difficult to formulate definitive tests of implied volatility performance in forecasting future spot volatility. Given this, there is conflicting evidence as to whether IV's from option pricing models provide superior forecasts to those obtained from spot data using appropriate econometric techniques (such as GARCH estimators). Some studies, such as Takezawa (1995), Canina and Figlewski (1993) and Stein (1989) find that the IV's tend to be poor forecasts of future realized volatility compared to econometric estimators. Among other reasons, this is attributed to implied volatilities 'over reacting' to volatility surprises. Other studies, such as Xu and Taylor (1994) find that implied volatilities are the best predictors of future volatility.

One difficulty in identifying the forecasting properties of implied volatilities concerns the sampling windows being used. On the IV side, there is the problem of how best to combine IV information. Implied volatilities change over time and numerous studies have demonstrated that averages, including weighted averages, of past IVs, have better performance than using just the previous IV, e.g, Schmalensee and Trippi (1978), Latane and Randleman (1976), Beckers (1981). IVs from at-the-money options are also found to be better predictors than deep out-of-the-money and, where applicable, deep in-the-money options. Short dated options and options written on extremely high or low volatility spot markets also have poor performance. There is also the question of how to combine the estimates from options on the same security, with the same time to expiration but different exercise prices. Another complication concerns the variable being forecasted. Is accuracy in pricing subsequent options the criteria, or is it the tracking of actual spot market volatility, or is it profitability of a trading strategy based on option mispricing, e.g., Galai (1977)? All these questions are complicated by non-simultaneity of option and spot quotes, as well as the problem of accounting for the bid/offer spread, e.g., Rubinstein (1985).

From these early results, the study of implied volatilities progressed to the detailed comparison of volatilities extracted from different options for the same commodity or security, concentrating on currencies and stocks. More precisely, at any point in time, a number of options with varying term to maturity and exercise price are available for the same spot commodity. The implied volatilities for the different option prices reveal a significant degree of heterogeneity. In particular, comparing implied volatilities for different exercise prices reveals a **volatility smile** or volatility smirk when currency option prices are used to calculate the IVs.[22] The IV is lowest for at-the-money options, with the IV having progressively higher values as the option price moves in-the-money or out-of-the-money, e.g., Heynen (1994), Dumitru (1999), Taylor and Xu (1994), Malz (2000).[23] This empirical result for currency options is stylized and the precise shape of the smile varies across time and currencies. For example, a smirk is sometimes observed where the IVs for at-the-money and in-the-money are approximately equal while the out-of-the-money options have higher IVs.

The IV smile of currency options does not carry over to stock and stock index options where a *volatility skew* is observed, e.g., Dumas et al. (1998). In a volatility skew, IV decreases as the strike price increases. Low strike price options, where the call is deep in-the-money and the put is deep out-of-the-money have the highest IVs. IV declines progressively reaching the lowest levels for high strike price options, where the put is deep in-the-money and the call is deep out-of-the-money. For both currency and stock options, differences in calculated IVs are also observed when different terms to expiration are used, with the shorter maturity options typically exhibiting higher IV's. All this variation in empirical IV has led to considerable theoretical analysis, attempting to explain the IV variation.

If the Black-Scholes model were perfect, then volatilities derived from option prices would be the same irrespective of the expiration date and exercise price of the option used to derive the IV. However, in practice, volatilities derived from in- at- and out-of-the-money options differ, sometimes the IV plot reveals a smile when implied volatility is plotted against moneyness of the option, e.g., Heynen (1994), and sometimes the plot reveals a smirk or a skew. A first possible step in dealing with observed IV behavior is to evaluate the *vega risk* of the position. For a position containing only one option, the vega risk can be approximated by taking the derivative, with respect to σ, of the appropriate Black-Scholes formula, e.g., Hull and White (1987), Malz (2000). This approach is discussed in Chapter 9. It is possible to incorporate vega risk into VaR calculations.

Another possible approach to accounting for observed IV behavior is to model the stochastic behavior of volatility directly. This involves introducing an additional stochastic process into the model. Unfortunately, the presence of two random variables, volatility and commodity price, significantly complicates the modelling process, e.g., Scott (1987, 1997), Hull and White (1987, 1988), Heston (1993), Ball and Roma (1994), Duan (1995), Ritchken and Trevor (1999). The basic problem of introducing volatility is captured by Ball and Roma: "Since volatility is not spanned by assets in the economy, the volatility risk may not be eliminated by arbitrage methods. Therefore, its market price of risk explicitly enters into the PDE." Hull and White (1987) were able to partially circumvent this problem by taking volatility to be uncorrelated with the commodity price, allowing the Black-Scholes price to be determined by integrating over the probability distribution of the average variance during the life of the option.

The Hull and White approach can be motivated using a power series expansion of the option price. Ball and Roma (1994) demonstrate the the Hull and White approach can be extended "to any security price process that is conditionally log-normal and for which the MGF of the average variance possesses a known analytical form". This follows because, if a tractable stochastic process for volatility has been specified, it is possible to solve for the expected variance by direct integration. This approach can be extended to demonstrate how stochastic volatility leads to a smile effect (Ball and Roma 1994, p.662). To see this, expand the option price in a univariate Taylor series, expanded in volatility around the expected value of the true variance of the underlying process. Ignoring terms higher than second order gives:

$$C* = C[E[\hat{\sigma}^2]] + \frac{1}{2}\frac{\partial^2 C[E[\hat{\sigma}^2]]}{\partial (\hat{\sigma}^2)^2} var[\hat{\sigma}^2] = C[\Omega] + \frac{1}{2}C'' \; var[\hat{\sigma}^2]$$

where C^* is the true option price that takes account of stochastic volatility, C is the Black-Scholes price, $E[\hat{\sigma}^2]$ $= \Omega$ is the expected value of the average variance, and var[$\hat{\sigma}^2$] is the variance of the average variance.

From this point, it is possible to adapt an argument similar to that in used in Sec. 2.1 to capture the cost of risk. Recall that IV is the implied volatility calculated from the Black-Scholes formula using the observed option price. While it is conventional to solve for IV using observed option prices, it is also possible to solve for the IV associated with the option price calculated from the pricing model that accounts for stochastic volatility. With this in mind, let $IV^* = \Omega + \omega$; the implied volatility calculated from the Black-Scholes model using the option prices determined by the stochastic volatility model is equal to the expected variance of the true process generating the commodity price *plus* any deviation of IV* from the expected variance of the process that accounts for stochastic variance. It follows that the first order Taylor series approximation gives:

$$C^* = C[IV] = C[\Omega + \omega] = C[\Omega] + \omega\frac{\partial C[\Omega]}{\partial(\hat{\sigma}^2)} \quad \rightarrow \quad \omega = \frac{1}{2}var[\hat{\sigma}^2]\,\frac{C''}{C'}$$

The relationship between the deviations of the implied volatilities calculated from Black-Scholes prices, on the same commodity, using the different prices available, can now be determined by directly evaluating the derivatives C'' and C'.

Ball and Roma give an approximation to the solution for ω as:

$$4\omega = \frac{var[\hat{\sigma}^2]}{\Omega}\left[-\{\frac{\Omega \; t^*}{4}\} + \frac{\ln[S/\{X \; e^{-rt^*}\}]^2}{\Omega \; t^*} - 1\right]$$

Solving this for at-the-money reveals $\omega < 0$. As Ball and Roma observe: "This is not surprising since the Black-Scholes option price is only everywhere concave in variance for at-the-money options and, hence, a straightforward application of Jensen's inequality implies that the stochastic volatility option is priced below the Black-Scholes counterpart evaluated at the mean of average variance." With this result the presence of a smile effect in the calculated IV using actual option prices can now be determined by observing that there are different ωs for the various exercise prices. Some caution is needed in applying these results due to the initial assumption that the randomness in volatility is uncorrelated with randomness in commodity prices. Failure of this assumption contributes to the volatility skew in equity option IVs.

Ultimately, only so much can be accomplished by dealing with σ, which is just one parameter from the underlying commodity or security price distribution. It is natural that attention has also focused on modeling or extracting the complete distribution for the underlying commodity price process. Improving pricing accuracy by better modeling the distribution of commodity prices has been of interested almost since Black-Scholes (1973). One potential approach along these lines is to empirically evaluate the commodity price distributions to assess which empirical assumption is most appropriate. For example, Bates (1996)and others examine the empirical fit of jump diffusion, stochastic volatility and mixed jump-stochastic volatility models. One inherent difficulty with this approach is that the distributional model which has the best goodness of fit, e.g., the skew-stable, may not be manageable theoretically, especially if the objective is a closed form solution.

Once it is recognized that there are only a few cases where closed forms can be derived, approximation methods gain appeal. Various approximation techniques, including Edgeworth expansions, Hermite-Fourier expansions Chiarella et al. (1999), Laguerre expansions (Dufresne 2000), and Gram-Charlier expansions, have been examined, e.g., Jarrow and Ruud (1982), Shimko (1994), Corrado and Su (1996). However, these methods are decidedly in the direction of numerical evaluation of option prices, e.g., using finite difference methods. While these methods can be used to accurately solve for option prices, the intuition and appeal of the closed form solution is lost.

Another related approach is to estimate the volatility surface directly, e.g., Derman and Kani (1994), Jackwerth and Rubinstein (1996), Avellaneda et al. (1997). To date, this approach has been found to be "no better than an ad hoc procedure that merely smooths Black-Scholes implied volatilities across exercise prices and times to expiration" (Dumas et al. 1998, p.2059).

8.3 Solving the Black-Scholes PDE

Possible Methods of Solving Black-Scholes

There are at least six methods available for deriving the Black-Scholes option pricing formula from the fundamental PDE. These methods are, in no particular order of importance: the original approach proposed by Black-Scholes, that involves making relevant substitutions into available solutions to the heat equation derived using Fourier series methods; the risk-neutral valuation approach of Cox-Ross, that involves solving the discounted expected value of the option's expiration date payoff function; evaluating the PDE as the limit of the binomial process; taking Laplace transforms; the Feynman-Kac approach, closely related to the risk-neutral valuation approach, that involves expressing the solution of the fundamental PDE as the expected value of a functional of Brownian motion; and, finally, by direct verification that the derivation of the Black-Scholes formula satisfies the PDE and boundary condition. While the last suggested method is consistent with 'proof by knowledge of the solution', it is also consistent with the ad hoc method, sometimes used in mathematical physics and engineering, of guessing a solution and then verifying that the proposed solution satisfies the PDE.

The method that has been found to have useful applications to a range of pricing situations, particularly where the solution to the pricing problem is not apparent, is the risk-neutral valuation method proposed by Cox and Ross. For the European call option on a non-dividend paying stock, this method involves directly evaluating the expectation:

$$C_t = e^{-rt^*} E\{\max\ [0,\ S_T - X]\}$$

$$= e^{-rt^*}\{\ E[S_T - X \mid S_T \geq X] + E[0 \mid S_T < X]\}$$

$$= e^{-rt^*}\{E[S_T - X \mid S_T \geq X]\}$$

$$= e^{-rt^*}\ \{E[S_T \mid S_T \geq X] - X\ Prob[S_T \geq X]\} \qquad (8.1)$$

where $t^* = T - t$, $E[\cdot]$ is the time t expectation taken with respect to the risk neutral density $Prob[\cdot]$ (see Appendix 3). Stoll and Whalley (1993) provide a worked solution for this problem.

As for the 'proof by knowledge of the solution' method of solving the Black-Scholes call option pricing formula, while this method is not conceptually satisfying, it is immensely practical in many situations. The method proceeds by examining the solutions for PDE problems with similar structure. As discussed in texts on PDEs, e.g., Berg and McGregor (1966), there is an elaborate classification scheme for PDE problems that facilitates this exercise. Having determined the solution to a closely related problem, a solution to the pricing problem is then guessed and derivatives of the pricing problem are taken to verify whether the PDE and boundary conditions are satisfied. This starts an iterative process that combines intuition and brute force to arrive at a solution. This method relies on the guidance of the existence and uniqueness theorems for PDEs from advanced mathematical analysis to determine when a solution will be available and if that solution is unique. Of course, a trivial variation of this method is to now the solution at the start and verify that it satisfies the PDE and boundary condition.

Yet another method to solve the Black-Scholes PDE involves applying the technique of Laplace transforms. Applied to the function $f[t]$, this transform is defined as:

$$L[\ f[t]\ ] = \int_0^\infty e^{-qt} f[t]\ dt$$

Solutions to the Laplace transform for a wide range of functions are available. After making an appropriate transformation of variables, the Laplace transform technique permits the fundamental PDE to be converted to an ordinary differential equation (ODE) that is easier to solve. In the Black-Scholes case, the Laplace transform solution involves the *erf[·]* function, that often appears in Laplace transformations involving probability distributions. The Laplace transform method is quite tedious and, without prior knowledge of the solution provided by Black-Scholes, it is not obvious how to proceed to the desired solution. Though widely used in engineering, the Laplace transform approach is not common is finance applications, e.g., Buser (1986).

The Original Black-Scholes Solution

The original notation in Black and Scholes (1973) differs somewhat from what is currently in common use. To retain the flavor of the Black-Scholes discussion, some of the original notation will be adopted. In Black-Scholes (1973), the notation used for the call price was *w* (to refer to warrants) and the PDE has the form:

$$\frac{\partial w}{\partial t} = rw - rS\ \frac{\partial w}{\partial S} - \frac{1}{2}\ v^2\ S^2 \frac{\partial^2 w}{\partial S^2}$$

where v^2 is the variance of the stock price. This PDE is also subject to the boundary condition applicable to the exercise value of the option on the maturity date *T*:

$$w(S,T) = S(T) - X \quad \text{for} \quad S(T) \geq X \quad \text{and} \quad w = 0 \ \text{ otherwise} \quad (S(T) < X)$$

At this point it is possible for analysis to proceed with 'proof by knowledge of the solution'. In other words, the Black-Scholes solution is "proved" by verifying that the Black-Scholes call option pricing function satisfies the PDE. This exercise will be carried out in Sec. 9.1. Restrictions imposed on the state space then make it possible to infer that the Black-Scholes solution is the unique solution to the PDE problem.

However, not knowing the solution to the PDE, Black and Scholes found it necessary to proceed by solving the boundary value problem directly. The method used is cumbersome and has since been superceded by other solution methodologies, such as direct evaluation of the discounted expectation problem used in risk neutral valuation. To better appreciate the more difficult and direct method of solving the PDE used by Black and Scholes, consider the following guide provided by Black and Scholes as to the method of finding a solution (p.643-4):

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

$$w(S,t) = e^{r(t-T)}\ y[(\frac{2}{v^2})\ (r - \frac{1}{2}v^2)\ [\ln\frac{S}{X} - (r - \frac{1}{2}v^2)(t - T)],$$

$$- (\frac{2}{v^2})(r - \frac{1}{2}v^2)^2(t - T)] \qquad (9)$$

With this substitution, the differential equation and boundary condition become:

$$\frac{\partial y}{\partial s} = \frac{\partial^2 y}{\partial u^2}$$

$$y(u,0) = 0 \qquad u < 0, \qquad\qquad y(u,0) = X\,[\exp\{\frac{u(\frac{v^2}{2})}{r - \frac{v^2}{2}}\} - 1] \qquad\qquad u \geq 0$$

This differential equation is the heat-transfer equation of physics, and its solution is given by Churchill (1963, p.155). In our notation, the solution is:

$$y(u,s) = \frac{1}{\sqrt{2\pi}} \int_{\frac{-u}{\sqrt{2s}}}^{\infty} X\,[\exp\{\frac{(u + q\sqrt{2s})(\frac{v^2}{2})}{(r - \frac{v^2}{2})}\} - 1]\, \exp\{\frac{-q^2}{2}\}\, dq$$

Substituting (this result) back into equation (9) from the Black-Scholes derivation, and simplifying provides the Black-Scholes formula.

The general method selected by Black-Scholes is to use a change of variable and scale to transform the fundamental PDE for the European call into a variation of the heat equation, e.g., Berg and McGregor (1966). This permits Fourier series results to be applied. In particular, Black and Scholes reference the solutions to the heat equation from Churchill (1963, p.155). This approach requires a complicated substitution to transform the Black-Scholes PDE into a form of the heat equation for which accessible solutions are available. Once this substitution is available, it is still not easy to derive a solution from the results in Churchill, and it is not immediately obvious that the form of the solution will involve density functions. Substantial effort is required to proceed from the solution derived from Churchill, given in terms of *y*, to the Black-Scholes call option pricing formula, that is given in terms of the call price *w*. It is the objective of this Section to provide that solution.

Churchill (1963, p.155) states a solution, derived using Fourier series methods, for the heat equation subject to a boundary condition of the kind applicable to Black-Scholes. The notation in Churchill differs from that used in Black and Scholes. Remembering that Black-Scholes did a change of variable transform from *w(S,t)* to *y(u,s)*, Churchill provides:

$$\textit{The solution to}: \quad \frac{\partial g(x,t)}{\partial x} = k\,\frac{\partial^2 g(x,t)}{\partial t^2} \qquad \textit{subject to } g(x,0) = f(x)$$

$$\textit{Is given by}:$$

$$g(x,t) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} f(x + 2\eta\sqrt{kt})\, e^{-\eta^2}\, d\eta$$

To match up this notation, change the dummy variable of integration to *q'*. Also observe that *k=1* in Black-Scholes and *g(x,t)* is written as *y(u,s)*. These substitutions produce:

$$y(u,s) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} y(u + 2q'\sqrt{s},\, 0)\, e^{\{-q'\}^2}\, dq'$$

It is now possible to exploit the properties of the boundary condition provided by the *max[·]* function to restrict the region of integration.

To do this, observe that $y(u,0) = 0$ *for* $u < 0$ and $y(u + 2q'\sqrt{s}, 0) = 0$ for $u = -2q'\sqrt{s}$. This can be solved for the dummy variable *q'* to get: $q' < -u/(2\sqrt{s})$. It follows that the lower limit of integration can be set to produce, in the Black-Scholes notation:

$$y(u,s) = \frac{1}{\sqrt{\pi}} \int_{-\frac{u}{2\sqrt{s}}}^{\infty} X \{\exp[\frac{(u + 2q' \sqrt{s})(\frac{v^2}{2})}{r - \frac{v^2}{2}} - 1] \exp[\{-q'\}^2] \, dq'$$

In effect, the Black-Scholes solution involves a change of variable from *w* to *y* to transform the PDE to a form that is comparable to a type for which solutions are more readily derived. This change of variable imposes specific solutions on the boundary condition associated with *y(u,0).* The properties of *y(u,0)* are used to restrict the range of integration and provide a solvable problem. Changing the variable of integration *q'* to be *(q /√2)* reproduces the Black-Scholes result.

What remains is to do as Black-Scholes suggests and "substitute back into (9)". The first step is to rewrite *y(u,s)* to get:

$$y(u,s) = \frac{1}{\sqrt{2\pi}} \int_{-\frac{u}{\sqrt{2s}}}^{\infty} X \exp[\frac{(u + q\sqrt{2s})(\frac{v^2}{2})}{(r - \frac{v^2}{2})}] \exp[\frac{-q^2}{2}] \, dq - \frac{X}{\sqrt{2\pi}} \int_{-\frac{u}{\sqrt{2s}}}^{\infty} \exp[-\frac{q^2}{2}] \, dq$$

Given this, the operation of substituting involves unbundling the change of variable. Observing that the *y(u,s)* specified by Black-Scholes in (9) provides the relevant definitions for *u* and *s*, the lower limit of integration can be solved as:

$$\frac{u}{\sqrt{2s}} = \frac{\frac{2}{v^2} (r - \frac{v^2}{2}) [\ln(\frac{S}{X}) - (r - \frac{v^2}{2}] (t - T)]}{\sqrt{\frac{4}{v^2} (r - \frac{v^2}{2})^2 (t - T)}} = \frac{\ln(\frac{S}{X}) + (r - \frac{v^2}{2})(t - T)}{v \sqrt{t - T}} = d_2$$

Using this result, the second term in *y(u,s)* becomes:

$$-\frac{X}{\sqrt{2\pi}} \int_{-d_2}^{\infty} \exp[\frac{-q^2}{2}] \, dq = -\frac{X}{\sqrt{2\pi}} \int_{-\infty}^{d_2} \exp[\frac{-\upsilon^2}{2}] \, d\upsilon = -X \, N[d_2]$$

The connection with the Black-Scholes formula is now apparent.

The first term of *y*(*u*,*s*) follows appropriately. Examining the numerator in the exponential, the following terms can be solved:

$$\frac{u \frac{v^2}{2}}{r - \frac{v^2}{2}} = \ln\frac{S}{X} + (r - \frac{v^2}{2})(t - T) = d_2 \, (v \sqrt{t - T})$$

$$\frac{\sqrt{2s} \frac{v^2}{2}}{r - \frac{v^2}{2}} = v \sqrt{(t - T)}$$

Using this the first term can now be written:

$$\frac{X}{\sqrt{2\pi}} \int_{-d_2}^{\infty} \exp[d_2\ \nu\sqrt{t\ -T}]\ \exp[q\ \nu\ \sqrt{t\ -\ T}]\ \exp[\frac{-q^2}{2}]\ dq$$

What follows is a considerable amount of manipulation. Using the following two results:

$$\exp[\frac{-q^2}{2}]\ \exp[q\nu\sqrt{t\ -\ T}]\ =\ \exp[-\frac{1}{2}(q\ -\ \nu\sqrt{t\ -T})^2]\ \exp[\frac{\nu^2(t\ -\ T)}{2}]$$

$$d_2\ \nu\ \sqrt{t\ -\ T}\ +\ \frac{\nu^2(t\ -\ T)}{2}\ =\ \ln\frac{S}{X}\ +\ r(t\ -\ T)$$

This produces for the first term:

$$\frac{X}{\sqrt{2\pi}} \int_{-d_2}^{\infty} \exp[\ln\frac{S}{X}\ +\ r(t\ -\ T)]\ \exp[-\frac{1}{2}(q\ -\ \nu\sqrt{t\ -\ T})^2]\ dq$$

$$=\ \frac{S}{\sqrt{2\pi}}\exp[r(t\ -\ T)] \int_{-d_2}^{\infty} \exp[-\frac{1}{2}\ (q\ -\ \nu\ \sqrt{t\ -\ T})^2]\ dq$$

What remains is to change the variable of integration:

$$q'\ =\ \nu\sqrt{t\ -\ T}\ -\ q \qquad \textit{this produces}:$$

$$\frac{S}{\sqrt{2\pi}}\ \exp[r(t\ -\ T)] \int_{-\infty}^{d_1} \exp[-\frac{1}{2}q']\ dq'\ =\ S\ \exp[r(t\ -\ T)]\ N[d_1]$$

Substituting the first and second terms into (9) from Black-Scholes reproduces the famous formula.

Risk-Neutral Valuation

For exposition purposes, it is instructive to consider two solutions to the risk-neutral valuation problem: the solution to the Black-Scholes option pricing problem, which assumes geometric Brownian motion for the state variable; and, the solution to the Bachelier option pricing problem, which assumes arithmetic Brownian motion for the state variable (Poitras 1998b). The solution to the risk-neutral Black-Scholes valuation problem for geometric Brownian motion is available in several sources, e.g., Stoll and Whalley (1995). The general risk neutral valuation problem, given in equation (8.1) starts by considering the valuation problem for a European call option on a non-dividend paying stock. In words, (8.1) indicates that the value of the call option depends only on the paths where $S_T \geq X$. The density associated with $S_T < X$, that includes the potentially negative price paths, does not directly enter the valuation. While there is some difference in the shape of the upper portions of the normal and log-normal distributions, this difference provides a basis for contrasting the performance of the Bachelier and Black-Scholes options. Given that neither assumption generally provides a particularly close fit to observed price distributions, it does not follow that lognormality will necessarily provide better pricing accuracy than normality in all situations. This issue must be addressed empirically.

Evaluation of (8.1) requires the integration of variables that follow the standard normal distribution. Two different specifications are needed to do this, one for the arithmetic Brownian motion and the other for geometric Brownian motion. For the case of arithmetic Brownian motion, over the time interval starting at t and ending at T, with $t^* = T - t$:

$$S_T = S_t + \alpha t^* + \sigma \sqrt{t^*}\, Z \qquad or \qquad Z = \frac{S_T - S_t - \alpha t^*}{\sigma \sqrt{t^*}}$$

where Z is $N(0,1)$. The standard normal specification associated with geometric Brownian motion is stated in Appendix III.

Evaluating (8.1) for the arithmetic Brownian motion case gives:

$$X\, Prob[S_T \geq X] = X\, N[\frac{S_t + \alpha t^* - X}{\sigma \sqrt{t^*}}]$$

$$E[S_T \,|S_T \geq X] = \int_{(X - S_t - \alpha t^*)/\sigma\sqrt{t^*}}^{+\infty} (S_t + \alpha t^* + \sigma\sqrt{t^*}\, Z)\, n[Z]\, dZ$$

$$= (S_t + \alpha t^*) \int_{-\infty}^{(S_t + \alpha t^* - X)/\sigma\sqrt{t^*}} n[Z]\, dZ + \sigma\sqrt{t^*} \int_{(X - S_t - \alpha t^*)/\sigma\sqrt{t^*}}^{+\infty} Z\, n[Z]\, dZ$$

$$= (S_t + \alpha t^*)\, N[\frac{S_t + \alpha t^* - X}{\sigma\sqrt{t^*}}] + \sigma\sqrt{t^*}\, n[\frac{S_t + \alpha t^* - X}{\sigma\sqrt{t^*}}]$$

Observing that risk neutrality requires that $(S_t + \alpha t^*) = S_t\, exp[r\, t^*]$, substituting these results into (8.1) gives the closed form solution for the Bachelier call option pricing formula. To verify that this is the absence-of-arbitrage consistent solution requires the derivatives of the fundamental PDE for the Bachelier call option formula stated in Proposition 8.1 to be evaluated and then substituted in the fundamental PDE for the Bachelier option. Determining that the PDE, together with the boundary and terminal conditions, is satisfied, ensures that absence of arbitrage is satisfied.

The risk neutral solution for the geometric Brownian motion proceeds much as with the arithmetic Brownian solution:

$$X\, Prob[S_T \geq X] = X\, N[\frac{\ln\left[\frac{S_t}{X}\right] + \mu\, t^*}{\sigma\sqrt{t^*}}]$$

$$E[S_T \,|S_T \geq X] = \int_{(\ln[X/S_t] - \mu t^*)/\sigma\sqrt{t^*}}^{+\infty} (S_t\, e^{\mu\, t^* + \sigma\sqrt{t^*}\, Z})\, n[Z]\, dZ$$

$$= \left[S_t\, e^{\mu\, t^* + \frac{\sigma^2\, t^*}{2}}\right] \int_{-\infty}^{\frac{(\ln[S_t/X] + \mu t^*}{\sigma\sqrt{t^*}} + \sigma\sqrt{t^*}} n[Z]\, dZ$$

$$= (S_t\, e^{\mu t^* + \frac{\sigma^2\, t^*}{2}})\, N[\frac{\ln[S_t/X] + \mu t^*}{\sigma\sqrt{t^*}} + \sigma\sqrt{t^*}]$$

From the discussion in Appendix III, the ***risk neutral*** valuation approach now permits the result that the riskfree interest rate *(r)* can be used for continuously compounded rate of return on the stock *(*α*)*. This produces $r = \alpha = \mu + \sigma^2/2$ to be used to make a substitution for μ $(= r - \sigma^2/2)$. Collecting all these results into (8.1) gives:

$$C(t) = e^{-rt^*}\{E[S(T) \mid S(T) \geq X] - X \; Prob[S_T \geq X]\} = e^{-rt^*}\{S(t) \; e^{rt^*} \; N[d_1] - X \; N[d_2]\}$$

Taking the discounting operator exp*{-rt*}* through gives the Black-Scholes formula.

8.4 Extending the Black-Scholes Model

Since the introduction of the basic Black-Scholes model, various extensions have been developed. This includes: permitting dividends on the stock; allowing early exercise, by valuing American options; assuming different diffusion processes for the stock price, such as the CEV process; allowing interest rates or volatility to be stochastic; applying the model to commodities other than stocks, including options that permit delivery of the spot commodity or future contracts; applying the model to warrants instead of exchange traded options; allowing options to have special features such as "look back" or multiple delivery specifications; introducing transactions and short selling costs; and, developing different analytical methodologies for solving the option pricing problem. This Section provides an overview of the work that has been done on some of these topics. However, except in a limited number cases, relaxing the basic Black-Scholes framework produces significant complications. For example, when American options are introduced, this produces ***path dependence*** in the option price effectively eliminating the possibility for a closed form solution to the valuation problem. Similarly, using other (potentially more realistic) diffusion processes than the log-normal again make it difficult to derive a closed form solution. For this reason, evaluation of option prices for more complicated situations typically involves application of techniques from numerical analysis.

Incorporating Dividends

Various methods are available for incorporating dividends into the Black-Scholes model. When the European assumption is retained and the timing and size of future dividend payments are known, only minor modifications are required to the formula.[24] The most direct method of incorporating dividends without undermining the basic Black-Scholes framework is to assume that the dividend payment is paid continuously, as a constant fraction or proportion (δ) of the value of the stock price, $D = \delta S$. It is also possible to be more general and assume that $\delta = \delta[S,\tau]$, but this complicates the derivation of the closed form. In the constant proportional dividend case, the Black-Scholes formula is altered to permit the stock price to be discounted using the dividend payment rate to produce the ***constant proportional dividend call option pricing*** formula:

$$C = S \exp\{-\delta t^*\} N[d_1] - X \exp\{-rt^*\} N[d_2]$$

where:

$$d_1 = \frac{\ln\{\frac{S}{X}\} + \{(r - \delta) + \frac{1}{2}\sigma^2\}t^*}{\sigma\sqrt{t^*}}$$

$$d_2 = d_1 - \sigma\sqrt{t^*}$$

This formula would be applicable to firms with stated dividend policies given in terms of ***yields*** as opposed to a fixed dollar payment. While not common, there are firms that follow such policies. More importantly, this formula can be readily extended to options on other types of securities, such as currencies.

To derive the constant proportional dividend call option formula requires modifying the return on the riskless hedge portfolio to be:

$$dV = [S - \frac{C}{C_S}]\, r\, dt - \delta S\, dt = \{S(r - \delta) - \frac{C}{C_S} r\}\, dt$$

In words, the net investment of funds in the riskless hedge portfolio must earn the riskless rate of interest adjusted for the dividend payments received. Because the dividend payment only enters into the return on the riskless hedge portfolio, the relevant partial differential equation becomes:

$$\frac{\partial C}{\partial t} = rC - (r - \delta)S\,\frac{\partial C}{\partial S} - \frac{1}{2}\frac{\partial^2 C}{\partial S^2}$$

It can be verified by direct differentiation that the constant proportional dividend call option pricing formula satisfies this PDE.

Incorporating known, discrete dividends follows in a similar fashion. Recalling the approach used in Sec. 6.1, when payment dates and amounts for dividends are known with certainty all that is required is to adjust the stock position in the riskless hedge portfolio by the ***appropriately*** discounted value of the dividends occurring between the purchase date and the expiration date. For a single dividend payment D_1 paid at time t^*_1 $(t^* > t^*_1 > 0)$, this leads to:[25]

$$C = [S - D_1 \exp\{-rt^*_1\}]\; N[d_1'] - X \exp\{-rt^*\}\; N[d_2']$$

where:

$$d_1' = \ln\{\frac{S - D_1 \exp\{-rt^*_1\}}{X \exp\{-rt^*\}}\} + \frac{1}{2}\sigma\sqrt{t^*}$$

$$d_2' = d_1' - \sigma\sqrt{t^*}$$

Similarly, incorporating multiple dividend payments occurring at times t^*_1, t^*_2, t^*_3 involves taking the discounted value of each dividend payment from the relevant payment dates, and subtracting these values from the current stock price. The implication of this result for option prices is that European options on dividend paying stocks will have lower call prices and higher put prices when compared with European options identical in X and t^* except that the stock does not pay a dividend. While seemingly an artificial method of introducing dividends, given the short maturities (9 months and less) for many exchange traded options, combined with the typically stable dividend payout patterns of many stocks, this approach can often provide reasonable approximations to observed option prices.[26]

Application of the Black-Scholes formula with Dividends

Consider the following information observed for a dividend paying stock on February 8, 1988 with call option expiring on March 25, 1988 (a leap year): S = \$51.70, X = \$52, r = 5.61%, t^* = 46/366 = 0.125683, σ = 0.1235. Assume that a quarterly dividend payment of \$1.50 is expected on March 15, 1988: t^*_1 = 36/366 = 0.09836. Substituting this value into the call option pricing formula adjusted for dividends gives: C = {51.7 - 1.5 exp{(-.0561)(0.09836)} $N[d_1']$ - 52 exp{(-.0561)(.125683)} $N[d_2']$. Evaluating $N[\cdot]$ gives $N[d_1']$ = 0.271 and $N[d_2']$ = 0.2546. This produces C = \$0.37. This result can be contrasted with the same parameters, ***excluding the dividend payment***. In this case, C = 51.7 $N[d_1]$ - 52 exp{(-.0561)(0.125683)} $N[d_2]$. Observing that d_1 = .290611 and d_2 = .248026 and evaluating gives C = \$0.937. Hence, this example demonstrates that dividends can have a significant impact on option pricing.

Given the formulae for incorporating known dividend payments, an important practical question concerns the extent to which these models capture actual option prices for dividend paying stocks, particularly for longer dated options such as warrants and LEAPS. This is currently an active research area, e.g., Lauterbach and Schultz (1990). In general, extending the Black-Scholes model to include uncertain dividend streams is not possible because this introduces a path dependent payoff. While it is possible to introduce a stochastic process for dividends, producing a model with two state variables, the resulting option pricing solutions are complicated. (Rabinovitch (1989) provides a workable solution for the related two state variable case with stochastic interest rates.) In addition, the stochastic process generating dividends is not typically the same as for variables such as interest rates and volatility.

Options on Futures and Forward Contracts

The techniques and issues associated with dividends arise naturally when the Black-Scholes framework is extended to options on futures contracts, e.g., Black (1976), Wolf (1982), Brenner, et al. (1985), Shastri and Tandon (1986), Ramaswamy and Sundaresan (1985), Hull (1989, Chap. 6). Starting from the Black-Scholes assumptions, the futures option pricing model can be derived most expediently by assuming that the commodity underlying the futures contract does not pay a carry return, this permits the relationship between spot and futures prices to be expressed as: $F(t) = S(t) \exp\{r t^*\}$.[27] In general, when a carry return is admitted: $F(t) = S(t) \exp\{bt^*\}$, where b is the **net** carry return, the continuous carry cost rate minus the continuous carry return rate. When the stock or spot commodity price is assumed to follow a lognormal diffusion: $dS = \alpha S\, dt + \sigma S\, dW$, *then* $dF = (\alpha - b)F\, dt + \sigma F\, dW$.[28] Given this, the riskless hedge portfolio used to derive the PDE for futures option contains one futures contract and $(\partial C/\partial F)^{-1}$ written options positions:

$$V = -\beta_F\, C \quad \Rightarrow \quad dV = [-\beta_F\, C]\; r\; dt \qquad where: \quad \beta_F = [\frac{\partial C}{\partial F}]^{-1} = \frac{1}{C_F}$$

Unlike stock (or spot commodity) options, the futures contract can be assumed not to involve a net investment of funds.[29] The only relevant cash flow is the income received on the written options position.

The derivation of the Black-Scholes formula for futures options, the fundamental PDE for a call option on a futures or forward contract can be derived as:

$$\frac{\partial C}{\partial t} = rC - \frac{1}{2}\frac{\partial^2 C}{\partial F^2}\,\sigma^2 F^2$$

where C is the price of a call option on a futures contract. This PDE differs substantively from that for a call option on the stock. This PDE can be compared with the fundamental PDE for a stock that pays a continuous dividend at rate δ given previously:

$$\frac{\partial C}{\partial t} = rC - (r - \delta)S\,\frac{\partial C}{\partial S} - \frac{1}{2}\frac{\partial^2 C}{\partial S^2}\sigma^2 S^2$$

Comparing the two PDEs reveals that, when $r = \delta$, there is an approximate equivalence relationship. From this is follows that the formula for pricing options on futures contracts can be inferred from the formula for the constant proportional dividend option pricing formula, under the condition that $r = \delta$:

$$C = \exp\{-rt^*\}\{F(t,T)\; N[d_1] - X\; N[d_2]\}$$

where:

$$d_1 = \frac{\ln\{\frac{F(t,T)}{X}\} + (\frac{\sigma^2}{2})t^*}{\sigma\sqrt{t^*}} \qquad d_2 = d_1 - \sigma\sqrt{t^*}$$

The validity of this formula depends on the noted assumptions about both the relationship between futures and spot prices as well as the constancy of the futures price volatility. Extensions of this model appear in a number of sources, e.g., Ramaswamy and Sundaresan (1985) and Brenner, et.al. (1985).

American Options

To this point, the extensions to Black-Scholes considered have permitted closed form solutions. Unfortunately, for many extensions of practical value, closed form solutions are not possible. The problem of path dependence has already been encountered in the analysis of introducing dividend payments and the American option feature. Because of the possibility for early exercise, the price of American puts as well as American calls on dividend paying stocks cannot be precisely determined ***for a finite expiration date***. Whether early exercise actually occurs will depend on the specific path that the stock price follows, hence the problem of path dependence. Given this, Roll (1977), Geske (1979) and Whaley (1981) show that it is possible to approximate the value of an American call by valuing hypothetical portfolios of European options that capture the American option payoff. Various analytical approximation methods have also been proposed, e.g., Brennan and Schwartz (1977), Geske and Johnson (1984), and Barone-Adesi and Whaley (1987). Tilley (1993) is an important recent contribution. Comparison of the American pricing models with the appropriately adjusted Black-Scholes formula, e.g., Shastri and Tandon (1986), reveals a significant, but not dramatic improvement in pricing accuracy.

Formal analysis of American options begins with McKean (1965) where the American option pricing problem is formulated as a solution to the heat equation subject to a free boundary condition. McKean demonstrated that the American option price can be determined but only by introducing a function that has come to be called the ***optimal stopping boundary***. Hence, the American option pricing problem can be formulated as an optimal stopping problem where the stopping boundary is determined by the benefits of early exercise. As a consequence, the price of an American option can be formally decomposed into the value of a European option plus the early exercise premium, producing a functional representation for Property 3. Since McKean, the optimal stopping representation has been developed considerably and other solution techniques such as variational inequalities have been introduced. Myneni (1992) provides an overview of these developments. One important implication of formalizing American option pricing as a free boundary problem occurs when the option is a perpetual. In this case, time does not directly enter the PDE problem and a closed form solution is available. This insight has been used to price other types of perpetuals such as the Russian option, e.g., Gerber and Shu (1994), Shepp and Shirayev (1993).

Recognizing that the American option price can be represented as the sum of the European price and the early exercise premium provides some intuitive insight. When the early exercise premium is zero, then the European option pricing formula can be used to value the American option. Because the right of early exercise combined with non-negative prices prevents the American option price from falling below $max[0, S(t) - X]$ for a call and $max[0, X - S(t)]$ for a put, early exercise cannot occur when the European option price is greater than these values, even when the early exercise premium is zero. It will be rational to sell the option instead of exercising. Because the European option achieves these values when $t=T$, it follows that one general factor involved in the early exercise decision is the time to expiration. However, even though the early exercise premium goes to zero as the time to expiration goes to zero, higher values of the time to expiration imply a larger time value for the European option. Recognizing that early exercise involves foregoing the time value on the option, the precise impact of time to expiration is not obvious.

Under what conditions will American options be exercised early? Rationality requires that the option be exercised only when the price received by selling the option is less than the value when exercised. Hence, one limited answer to this question is obvious: only ***in-the-money*** options will be exercised. Another early exercise Property that has already been provided is: American calls on non-dividend paying stocks will ***not*** be exercised early. This was an implication of Property 10 in Sec. 7.1. The European option price provides a lower bound form the American price. The arbitrage supporting the European call price prevents that price from falling below $\{S(t) - X\}$. Hence, early exercise for ***calls*** has something to do with dividends or, in the case of commodities other than stocks, with carry returns.[30]

Take the case of a stock with one dividend paying remaining prior to the expiration date on the option, the arbitrage support for the European call price requires the stock price to be adjusted for the dividend payment:

$$C_A[S,\tau,X] \geq C[S,\tau,X] \geq Max[0, S(t) - (D\ PV[r,\tau_1]) - X\ PV[r,\tau]]$$

The requirement of adjusting the stock price for the receipt of the dividend payment creates a situation where the lower bound to the American call price provided by the European call price is ***not*** necessarily greater than or equal to $S(t)$ - X. For a given τ and r, the larger is the dividend payment, the farther below $S(t) - X$ is the European bound on the American. For a given r and D, the same result applied for smaller τ. The implication is that short dated American options on stocks expecting a sizable dividend will tend to be exercised early. The case of early exercise for continuous dividend payments is discussed in Sec. 8.4.

The early exercise incentive for an American call on a dividend paying stock is even more complicated than the European lower bound indicates. As discussed in Sec. 7.1, the European lower bound is derived by comparing the expiration date returns for a European call with a portfolio that combines a long stock position financed by a borrowing with maturity value of X. When there is dividend payments on the stock, the cost of the purchasing the portfolio is reduced by the appropriately discounted value of the dividends received. The American option provides the flexibility to exercise the call just prior to the ex-dividend date, thereby receiving the dividend and avoiding the ***expected*** fall in the stock price typically occurring on the ex-dividend date. This result can be intuitively seen by comparing the European lower bound on the ex-dividend date, $Max\ [0, S(t+1) - X\ PV[r,\tau+1]]$, with the value on the previous cum-dividend date, $Max\ [0, S(t) - (D\ PV[r,\tau_1]) - X\ PV[r,\tau]]$.

Early exercise for puts on nondividend paying stocks is different than the case for calls. Consider the extension of Property 10 to puts:

$$P_A[S,\tau,X] \geq P[S,\tau,X] \geq Max[0, X - S(t)] \geq Max[0, X\ PV[r,\tau] - S(t)]$$

Unlike calls, the impact of $PV[\cdot]$ acts in the opposite direction on X making the possibility of early exercise for puts on nondividend paying stocks (or commodities where the carry return is less than r) a likely possibility.

Because the conditions developed from Property 10 are only lower bounds, American put prices will typically be higher than the lower bound. The intuition for early exercise follows from recognizing that for the stock price paths going close to zero, the put will be deep-in-the-money. This implies that $N[-d] \rightarrow 1$ and $P \rightarrow (PV[\cdot]\ X - S) \leq X - S$. Instead of selling the put for $(PV[\cdot]\ X - S)$, the American put holder will exercise the put and receive $X - S$. In practice, exercise of a put requires that the stock be delivered in exchange for X. Hence, early exercise will be done by traders holding stock in combination with a put. The essential point to recognize is that Property 10 of Sec. 7.1 applied to puts provides an arbitrage strategy that bounds the European option price. Because this bound is ***below*** the exercise value, holders of American options will choose to exercise the put rather than to sell the option at a lower price. Hence, deep in-the-money American puts will tend to be exercised early.

In addition to establishing early exercise of the American put using the distribution-free arbitrage bounds, the optimal stopping boundary for American puts can be used to motivate early exercise activity related to deep in-the-money puts. Puts are in-the-money because the stock price has fallen below the exercise price. At some point it is no longer profitable to hold the put because the possibility of further profits due to lower stock prices is more than offset by the loss of interest income on the exercise value, together with the interest opportunity cost of holding a non-dividend paying stock (spot) position. This leads to the simplest optimal stopping boundary: at any time between purchase and expiration, exercise the American put on a non-dividend paying stock when the loss of interest income on the exercise value plus the opportunity cost of holding the spot position exceeds the difference between the current put price and the exercise value.

The need to adjust for the difference between the put price and the exercise value is that selling the put is an alternative to early exercise. Where a stock and a put position are involved, as is required for exercise, transactions costs will impact the exercise decision. Selling the put will incur transactions costs in both option and stock markets, while exercise only involves delivery. When the put is exercised, there is some loss because the time premium is lost. This premium reflects that maximum future gain from the ***expected*** drop in stock prices between the current date and the expiration date. When the option goes deep-in-the-money or ***when*** $\tau \rightarrow \mathbf{0}$ and the option is close to expiration, this additional premium will be near zero. This brings out another important feature of early exercise: put options that are in-the-money and close to expiration may get exercised early.

Option Valuation with Alternative Diffusion Processes

Empirically, there is little support for the hypothesis that financial prices conform to the log-normal assumption that is essential to deriving the Black-Scholes option pricing formula. However, log-normality does have the desirable feature that non-negative values are not admitted, consistent with financial prices, and the deviation from observed prices is not dramatic. There are better distributional fits to the data, but log-normality does typically provide a useful first approximation. Unfortunately, making the distributional assumption more realistic also undermines the ability to derive a closed form option pricing formula. This issue is made even more complicated by a lack of agreement over which distributional assumption has the best goodness of fit with observed prices. In addition, it is not clear how much of an increase in pricing accuracy can be gained by over the Black-Scholes solution if more complexity is introduced.

Closed form solutions are available for a limited number of other distributional assumptions. The simplest solution occurs when it is assumed that prices follow ***arithmetic Brownian motion***: $dS = \alpha\ dt + \sigma\ dW$. Progressively more complicated solutions can be obtained for: a stochastic process which is a combination of a diffusion process and a Poisson jump process (Merton 1976); the constant-elasticity of variance process (Cox and Ross 1976); and the Bessel process (Geman and Yor 1993). Solutions for related processes, such as the Brownian bridge (Ball and Torous 1983) are also available. Jarrow and Rudd (1983) collects many relevant results.

In the arithmetic Brownian case, it is the change in prices follows an arithmetic Brownian motion, not the rate of return as in the log-normal case. Reasons for selecting geometric over arithmetic Brownian motion were advanced at least as early as Samuelson (1965) and some of the studies in Cootner (1964). In reviewing previous objections, Goldenberg (1991) recognizes three of practical importance: (i) a normal process admits the possibility of negative values, a result that is seemingly inappropriate when a security price is the relevant state variable; (ii) for a sufficiently large time to expiration, the value of an option based on arithmetic Brownian motion exceeds the underlying security price; and, (iii) as a risk-neutral process, arithmetic Brownian motion without drift implies a zero interest rate. Taken together, these three objections are relevant only to an unrestricted, risk-neutral "arithmetic Brownian motion" that is defined to have a zero drift. As such, some objections to arithmetic Brownian motion are ***semantic***, avoidable if the process is appropriately specified.

Smith (1976) defines arithmetic Brownian motion to be driftless and provides an option pricing formula that is attributed to Bachelier (1900) and is subject to all of the three objections. Smith (p.48) argues that objection (ii) for the formula is due to the possibility of negative sample paths, though this objection can also be avoided by imposing an appropriate drift. Goldenberg (1991) reproduces the Smith-Bachelier formula and proceeds to alter the pricing problem by replacing the unrestricted driftless process with an arithmetic Brownian motion that is absorbed at zero. The resulting option pricing formula avoids the first two objections. The third objection is addressed by setting the drift of the arithmetic Brownian process equal to the riskless interest rate times the security price, consistent with an OU process. Using this framework, Goldenberg generalizes an option pricing result in Cox and Ross (1976) to allow for changing variances and interest rates. With the use of appropriate transformations for time and scale, Goldenberg argues that a wide range of European option pricing problems involving diffusion prices can be handled using the absorbed-at-zero arithmetic Brownian motion approach.

The Bachelier option prices derived here use the same perfect market and continuous trading assumptions as in Black and Scholes (1973). Bachelier options for three types of underlying securities are considered: a non-dividend paying asset; an asset with proportional dividend payments; and, a futures contract. For the non-dividend paying asset, the security price is assumed to follow an unrestricted arithmetic Brownian motion:

$$dS = \alpha\ dt + \sigma\ dW$$

At this point, the drift α and volatility σ are assumed to be constant, though this assumption will be demonstrated as incompatible with absence-of-arbitrage. Exploiting the riskless hedge portfolio construction leads to the following PDE associated with the Bachelier option:

$$\frac{\partial C}{\partial t} = rC - \frac{\partial C}{\partial S} rS - \frac{1}{2} \sigma^2 \frac{\partial^2 C}{\partial S^2}$$

that is subject to the terminal boundary condition $C[S_T] = max[0, S_T - X]$. Solving the general valuation problem leads to the following:

Proposition 8.1: **Absence-of-Arbitrage Formula for the Bachelier Option**[31]

Assuming perfect markets and continuous trading, if the non-dividend paying security price follows the unrestricted arithmetic Brownian motion, then the solution to the absence-of-arbitrage valuation problem is given by:

$$C_G(S, t^*; r, \sigma, X) = e^{-rt^*}\{(S_t e^{rt^*} - X) N[g] + \sigma \sqrt{t^*} n[g]\}$$

where:

$$g = \frac{S_t e^{rt^*} - X}{\sigma \sqrt{t^*}}$$

C_G is the price of the general Bachelier call option, and *N[g]* and *n[g]* represent the cumulative normal density and normal probability function, respectively, evaluated at g.

Wilcox (1990) adapts this form of the Bachelier option to provide a solution to the spread option pricing problem.

It is possible to generalize the analysis to include constant proportional dividend payments (D) of the form: $D = \delta S$. Following market convention, the option contract is assumed to be **not** dividend-payout protected. To derive a Bachelier option that both satisfies the PDE associated the riskless hedge portfolio **and** incorporates proportional dividend payments, it is sufficient to restate the arithmetic Brownian motion as an OU process of the form:

$$dS = (r - \delta)S\, dt + \sigma_V\, dW$$

The associated PDE for the riskless hedge portfolio problem has the form:

$$\frac{\partial C}{\partial t} = rC - (r - \delta)S \frac{\partial C}{\partial S} - \frac{1}{2} \sigma_v^2 \frac{\partial^2 C}{\partial S^2}$$

Even though the drift does not enter the PDE directly, it does alter both the variance of the process and, as reflected in Proposition I, the expected future spot price. It follows that:

Proposition 8.2: **Absence-of-Arbitrage Valuation Formula for the Bachelier Option, with Dividends**

Assuming perfect markets and continuous trading, if the proportional dividend paying security price follows the arithmetic Brownian motion with continuous proportional dividends, then the absence-of-arbitrage solution to the valuation problem is given by:

$$C_R(S, t^*; r, \sigma, X) = (S_t e^{-\delta t^*} - Xe^{-rt*}) N[h] + V n[h]$$

where:

$$h = \frac{S_t\, e^{-\delta t^*} - X e^{-r t^*}}{V} \qquad V = \sigma_V \sqrt{\left\{\frac{e^{-2\delta t^*} - e^{-2 r t^*}}{2(r - \delta)}\right\}}$$

and *N[h]* and *n[h]* represent the cumulative normal density and normal probability function, respectively, evaluated at *h*.

The proof of Proposition 8.2 involves direct differentiation of the option pricing formula to verify that this solution does satisfy the PDE for the riskless hedge portfolio problem.

Despite certain similarity, there are substantive differences in option prices derived using geometric and arithmetic Brownian motion. For example, the PDE for the riskless hedge portfolio associated with options on futures contracts for arithmetic Brownian motion involves setting $r = \delta$ in the SDE for the OU process:

$$\frac{\partial C}{\partial t} = rC - \frac{1}{2}\sigma^2 \frac{\partial^2 C}{\partial F^2}$$

where F is the price for the relevant futures contract. A similar result holds for the geometric case. However, when it comes to solving for the option price formula, the simplification available for the geometric case (setting $r = \delta$ in the option price formula for stocks with proportional dividends) is not available in the arithmetic case due to a singularity in the Proposition 8.2 solution at $r=\delta$. The appropriate solution is:

Proposition 8.3: **Absence-of-Arbitrage Valuation for the Bachelier Futures Option**

Assuming perfect markets and continuous trading, if the futures price follows an arithmetic Brownian motion, then the absence-of-arbitrage solution to the European call option valuation problem is given by:

$$C_F(F,\ t^*;\ r,\ \sigma,\ X) = e^{-r t^*}\{(F_t - X)\ N[k] + \sigma \sqrt{t^*}\ n[k]\}$$

where:

$$k = \frac{F_t - X}{\sigma \sqrt{t^*}}$$

C_F is the price of the Bachelier futures call option, and *N[K]* and *n[k]* represent the cumulative normal density and normal probability function, respectively, evaluated at *k*.

Proposition 8.3 is useful both for pricing options on futures spreads and for comparing the normal and log-normal prices of exchange options.

The question of whether arithmetic Brownian motion is an appropriate stochastic process for asset pricing is complicated. In any event, conventional wisdom maintains that an important limitation of the OU and arithmetic Brownian processes is that the state variable has a non-zero probability of taking on negative values. Because prices are often the state variables in continuous time models, some method of bounding the variable at zero is needed. The log-normal process, geometric Brownian motion, has a natural boundary at zero. Another possible method of imposing non-negative state variable values is to impose an ***absorbing barrier*** at zero. A solution to the option pricing problem using absorbed Brownian motion was initially presented by Cox and Ross (1976) and later extended by Goldenberg (1991).

Converting an unrestricted arithmetic Brownian motion by imposing an absorbing barrier at zero on the/ process is not as difficult as it might seem.[32] All those paths that hit zero end and, from that point, do not contribute to the ensemble of paths that combine to determine the distribution at any future point in time. The solution to this option pricing problem was given by Cox and Ross (1976, p.162-3):

Proposition 8.4: Absence-of-Arbitrage Valuation for the Absorbed Brownian Option

Assuming perfect markets and continuous trading, if the proportional dividend paying security price follows the arithmetic Brownian motion of Proposition 8.3 subject to an absorbing barrier at zero, then the absence-of-arbitrage solution to the valuation problem is given by:

$$C_{Ab}(S,\ t^*;\ r,\ \sigma,\ X) = (S_t\ e^{-\delta t^*} - Xe^{-rt*})\ N[h_1] + (S_t\ e^{-\delta t^*} + Xe^{-rt^*})\ N[h_2] + V\ (n[h_1] - n[h_2])$$

where:

$$h_1 = \frac{S_t\ e^{-\delta t^*} - Xe^{-rt^*}}{V} \qquad V = \sigma_V \sqrt{\{\frac{e^{-2\delta t^*} - e^{-2rt^*}}{2(r - \delta)}\}} \qquad h_2 = \frac{-S_t\ e^{-\delta t^*} - Xe^{-rt^*}}{V}$$

and *N[h]* and *n[h]* represent the cumulative normal density and normal probability function, respectively, evaluated at *h*.

Extensions of the absorbed arithmetic Brownian motion, e.g., to allow for volatility being a deterministic function of time, can be found in Goldenberg (1991).

In the discussion of implied volatility, it was observed that it was possible to construct pricing results using two stochastic processes, one for the commodity price and another for volatility. Hull and White (1987), Heston (1993), Scott (1997) have demonstrated that it is possible to obtain closed form solutions for the case of two random variables. However, analytical solutions are substantively simplified in cases where the two random processes are assumed to be uncorrelated. This approach is not restricted to volatility. For example, in addition to random exchange rates, currency option pricing can also model the randomness in the foreign and domestic interest rates, pricing the option written on three stochastic processes. Other examples would include: introducing a stochastic dividend yield to price a stock index option; using a stochastic short interest rate process to price a Tbond option; and, using a stochastic convenience yield to price an option on oil futures. In general, it is only possible to achieve closed form solutions for a very small class of stochastic processes.

Rabinovitch (1983) is an example of a case where the two stochastic processes are chosen carefully enough to permit a closed form solution to be derived. As in Black-Scholes, Rabinovitch assumes that the stock price follow a log-normal diffusion: $dS = \alpha\ S\ dt + \sigma\ S\ dW_S$. In addition, there is a single default-free interest rate r that is also permitted to be stochastic, to follow a mean-reverting OU (Ornstein-Uhlenbeck) process of the form:

$$dr = q(m - r)\ dt + v\ dW_r$$

where $q(m - r)$ is the instantaneous expected change in the short-term rate, v^2 is the instantaneous variance for the interest rate. The parameter m is the unconditional expected interest rate to which the current rate, r, reverts to at speed proportional to q. This type of process is also (descriptively) called a mean-reverting process.[33] The presence of the two Wiener processes, dW_r and dW_S requires specification of a contemporaneous covariance, $\{dW_r\ dW_S\} = \rho\ dt$. From the Markov construction, lagged correlations are zero. Within this framework, it is possible to derive a closed-form expression for the call price that is similar in form to Black-Scholes. However, in order to do this, it is necessary to convert the information in the stochastic interest rate into a formula for the discount bond price. The resulting solutions, while expressible in the same form as Black-Scholes, are significantly more complicated.

8.4 Application: Foreign Currency Options

The Institutional Background

– Lifetime – High	Low	Month	Open	— Daily — High	Low	Settle	Chg	Open Interest
Currency								
Australian Dollar (IMM)								
A$100,000 US$ per A$; 0.0001 = $10 per contract								
.7455	.6956	Dec94	.7411	.7419	.7400	.7411	......	6,220
Est.Vol. 220,				Prev.Vol. 802,			Prev.OpenInt.	6,275
British Pound (IMM)								
£62,500 US$ per £; 0.0002 = $12.50 per contract								
1.6392	1.4400	Dec94	1.6190	1.6376	1.6176	1.6346	+.0128	45,442
1.6440	1.4528	Mar95	1.6190	1.6360	1.6190	1.6330	+.0124	529
Est.Vol. 23,694,				Prev.Vol. 24,249,			Prev.OpenInt.	45,980
Canadian Dollar (IMM)								
$100,000 US$ per C$; 0.0001 = $10 per contract								
.7680	.7038	Dec94	.7394	.7398	.7386	.7392	-.0011	33,877
.7673	.7020	Mar95	.7395	.7395	.7384	.7392	-.0011	1,717
.7600	.6990	Jun95	.7386	.7386	.7380	.7385	-.0011	821
.7438	.6965	Sep95	.7373	.7373	.7370	.7373	-.0011	692
Est.Vol. 2,582,				Prev.Vol. 2,701,			Prev.OpenInt.	37,170
French Franc (IMM)								
500,000 francs US$ per franc; 0.0001 = $5 per contract								
.1966	.1684	Dec94	.1940	.1942	.1941	.1941	+.0009	1,290
Est.Vol. N.A.,				Prev.Vol. 34,			Prev.OpenInt.	1,302
German Mark (IMM)								
125,000 marks US$ per mark; 0.0001 = $12.50 per contract								
.6731	.5590	Dec94	.6615	.6662	.6612	.6648	+.0023	88,871
.6745	.5797	Mar95	.6630	.6673	.6628	.6659	+.0022	5,383
.6747	.5980	Jun95	.6655	.6680	.6655	.6676	+.0023	1,249
.6770	.6259	Sep95	.6692	.6692	.6692	.6692	+.0023	116
Est.Vol. 26,920,				Prev.Vol. 87,282,			Prev.OpenInt.	95,619
Japanese Yen (IMM)								
12.5 million yen US$ per yen (scaled .00); 0.0001 = $12.50 per contract								
1.0490	.9525	Dec94	1.0305	1.0377	1.0296	1.0358	+.0045	61,422
1.0560	.9690	Mar95	1.0405	1.0458	1.0403	1.0441	+.0045	7,085
1.0670	.9776	Jun95	1.0541	1.0541	1.0541	1.0541	+.0045	718
1.0775	1.0197	Sep95	1.0637	1.0637	1.0637	1.0637	+.0045	180
Est.Vol. 16,843,				Prev.Vol. 31,511,			Prev.OpenInt.	69,473
Swiss Franc (IMM)								
125,000 francs US$ per franc; 0.0001 = $12.50 per contract								
.8108	.6852	Dec94	.7931	.7995	.7931	.7971	+.0027	41,588
.8136	.7085	Mar95	.7976	.8023	.7976	.8004	+.0027	2,066
.8165	.7127	Jun95	.8045	.8062	.8045	.8045	+.0027	168
Est.Vol. 16,486,				Prev.Vol. 33,554,			Prev.OpenInt.	43,828
U.S. Dollar Index (FINEX)								
1000 x index points and US¢; 0.01 = $10 per contract								
99.27	84.95	Dec94	86.23	86.23	85.62	85.78	-.35	6,846
96.96	85.39	Mar95	86.31	86.31	85.84	85.93	-.35	3,086
Est.Vol. 1,500,				Prev.Vol. 5,087,			Prev.OpenInt.	9,967

A listing of the most important exchange traded foreign currency option contracts in the US given in Table 8.3. Unlike exchange traded options for most other commodities, the major US exchange for trading currency options was, for many years, not also a major commodity futures exchange. This situation has now changed and comparison of the volume and open interest available for the currency options reveals that the futures options on the IMM division of the CME now is well in excess of the volume and option interest on the Philadelphia Stock Exchange (PHLX) which was previously the most important exchange trading currency option contracts. This is consistent with the usual result that the futures exchange trading the highest volume of options contracts significant to currency options trading than the IMM, the most important currency futures exchange. The PHLX was something of an innovator in options trading practices. In addition to the offering American currency futures options, the PHLX also offers European spot currency options.

The Currency Option Pricing Formula

Derivation of the pricing formula for foreign currency options involves an adaption of the constant proportional dividend version of the Black-Scholes model. It is assumed that the spot exchange rate follows a log-normal diffusion, where the domestic interest rate, r, and the foreign interest rate, r^*, together with the instantaneous standard deviation σ of the spot exchange rate are all constants. This leads to the following formula for a European currency call option on spot exchange:

$$C = S \exp\{-r^* t^*\} N[d_1] - X \exp\{-rt^*\} N[d_2]$$

where:

$$d_1 = \frac{\ln\{\frac{S}{X}\} + (r - r^* + \frac{1}{2}\sigma^2)\, t^*}{\sigma \sqrt{t^*}}$$

$$d_2 = d_1 - \sigma \sqrt{t^*}$$

As with previous call option pricing results, the formula for currency put options is obtained by substitution into the put-call parity condition for currency options:

$$P = X \exp\{-rt^*\}\ N[-d_2] - S \exp\{-r^* t^*\}\ N[-d_1]$$

where:

$$-d_1 = \frac{\ln\{\frac{X}{S}\} + (r^* - r - \frac{1}{2}\sigma^2)\ t^*}{\sigma \sqrt{t^*}}$$

$$-d_2 = -d_1 + \sigma \sqrt{t^*}$$

Close examination of the put and call pricing formulae reveals a fundamental feature of currency options: the right to put currency *A* for currency *B* at *X* is identical to the right to call currency *B* in exchange for currency *A* at *X*. Unlike other commodities that involve the exchange of some commodity for money, currency options involve the exchange of two monies with the numeraire being arbitrary.

While the European currency option pricing formula can be derived by following much the same approach as for the Black-Scholes call option on a stock with a constant proportional dividend, because the single source of randomness is now a spot exchange rate, the conceptualization of the hedge portfolio is somewhat different. The riskless hedge portfolio requires an investment in a pure discount (zero coupon) foreign bond that matures to a maturity value of one unit of foreign currency. In foreign currency terms, the continuous time representation for the price of this bond is: $\exp\{-r^*t^*\} = PV^*[r^*,\tau]$. The value of this bond is then converted to a domestic currency value by multiplying by the spot exchange rate to get: $S(t)\ \exp\{-r^*t^*\}$. As currency options are typically priced as though the domestic bond is US, this implies that the exchange rate is measured in US direct terms. The objective of the riskless hedge portfolio is to protect (hedge) the domestic currency value of the foreign bond position.

The riskless hedge portfolio is now constructed by writing β currency call options to protect the domestic currency value of the foreign bond against changes in the spot exchange rate. However, in this case, β has a slightly different interpretation than for the option on the non-dividend paying stock:

$$\frac{\partial V}{\partial S} = e^{-r^* t^*} - \frac{\partial C}{\partial S} = 0 \qquad \Rightarrow \qquad \beta = \frac{e^{-r^* t^*}}{\frac{\partial C}{\partial S}} = \{N[d_1]\}^{-1}$$

More precisely: $V(t) = S(t)\ \exp\{-r^*t^*) - \beta C$ where $S(t)$ is the spot exchange rate at time t. Much as in the constant proportional dividend case, the riskless return associated with the net investment in the portfolio must now be adjusted to reflect the return earned on the foreign bond:

$$dV = \{S(t)\ e^{-r^* t^*} - \beta C\}\ r\ dt - \{S(t)\ e^{-r^* t^*}\}\ r^*\ dt$$

$$= S(t)\ e^{-r^* t^*}\ (r - r^*)\ dt - \beta C\ r\ dt$$

Following the usual procedure of evaluating Ito's Lemma:

$$dV = e^{-r^* t^*}\ dS - \frac{e^{-r^* t^*}}{\frac{\partial C}{\partial S}}\ dC = e^{-r^* t^*}\ [dS - \frac{1}{\frac{\partial C}{\partial S}}\ dC] = e^{-r^* t^*}\ [dS - \frac{dC}{C_S}]$$

$$dS = \alpha S\ dt + \sigma S dW \qquad dC = [C_t + C_S \alpha S + \frac{1}{2} C_{SS}\ \sigma^2\ S^2]\ dt + C_S \sigma S\ dW$$

Equating the two conditions for dV, canceling and manipulating where appropriate produces:

$$C_t = rC - S\,(r - r^*)\,C_S - \frac{1}{2}\,\sigma^2\,S^2\,C_{SS}$$

This is the fundamental PDE for a European spot currency call option.

The original presentation of the currency call option formula was by Garman and Kohlhagen (1983), Grabbe (1983) and Biger and Hull (1983). Though the formula presented was the same, the derivations provided in these sources did not follow the constant proportional dividend approach. For example, Grabbe (1983) shows that a one unit long position in the domestic bond can be hedged with $(-\alpha/\beta)$ units of the foreign bond combined with $(-1/\beta)$ written call positions, where α and β are the appropriate partial derivatives. In effect, by creating a self-financing riskless hedge portfolio, Grabbe utilized the concept of an arbitrage portfolio. Grabbe also provides a generalization of the formula to admit deterministic changes in the volatility over time. Garman and Kohlhagen also use a different approach relying on the equality of risk-adjusted excess returns in an arbitrage free economy. Chiang and Okunev (1993) demonstrate that the formula can be generalized to permit the foreign and domestic bond prices to follow Brownian bridge processes.

The extension to options on currency futures and forwards follows by using the continuous time version of covered interest parity. Recall from Sec. 4.2 that:

$$F(t,T) = \frac{1 + r(t,T)}{1 + r^*(t,T)}\,S(t) = \frac{1 + rt^*}{1 + r^*t^*}\,S(t)$$

$$\Rightarrow \qquad F(t,T) = e^{rt*}\,e^{-r^*t^*}\,S(t) = e^{(r - r^*)\,t^*}\,S(t) \qquad \Rightarrow \qquad S(t) = e^{(r^* - r)\,t^*}\,F(t,T)$$

Substituting for S in the spot currency call option formula produces:

$$C = e^{-rt^*}\,\{F(t,T)\,N[d_1] - X\,N[d_2]\}$$

$$where: \quad d_1 = \frac{\ln[F(t,T)/X] + \frac{1}{2}\sigma^2\,t^*}{\sigma\,\sqrt{t^*}} \qquad d_2 = d_1 - \sigma\,\sqrt{t^*}$$

Comparison of this formula with the general formula for options on futures contracts reveals that the two results are the same, confirming the application of the general model in the specific case of options on currency futures and forward contracts.

Put-Call Parity and Early Exercise for Currency Options

Results for currency put options follow from put-call parity arbitrage condition for European Foreign Currency Options. For spot currencies the condition is:

$$P[S,\tau,X] = C[S,\tau,X] - S(t)\,PV^*[r^*,\tau] + X\,PV[r,\tau]$$

Using the CIP condition and substituting for S, gives the put-call parity condition for futures contracts:

$$P[S,\tau,X] = C[S,\tau,X] + \{(X - F(t,T))\,PV[r,\tau]\}$$

The put-call parity result for spot currency options can be derived in much the same fashion as put-call parity for non-dividend paying stocks given in Sec. 7.3.

While there are a number of different possible portfolio pairings that could be compared for illustrating the arbitrage, for present purposes the two relevant portfolios are: the domestic portfolio, a long (short) foreign currency call with exercise price X and τ days to maturity combined with a long (short) position in a domestic zero coupon bond with maturity value of X; and, the foreign portfolio, a long (short) foreign currency put-- with the same terms as the call-- combined with long (short) foreign zero discount bond yielding r^* with maturity value of one unit of foreign currency and term to maturity of τ.[34] When $S(T) \leq X$, then the domestic portfolio has a value equal to the maturity value of the bonds X, because the call expires worthless. The foreign portfolio will have maturity value of $S(T)$ plus $(X - S(T))$ that gives a value of X. Similarly, when $S(T) > X$, the domestic portfolio will be worth $S(T)$ because the call will be worth $(S(T) - X)$ and the domestic bonds X. The foreign portfolio will be worth the maturing value of the bonds $S(T)$, with the put expiring worthless.

The put-call parity results have significant practical implications. For example, the notion of covered interest arbitrage discussed in Sec.5.2 required that securities which were identical in all respects except currency of denomination will have the same fully covered returns. In the put-call parity case, another somewhat more complicated equivalence result is provided. More precisely, a foreign bond portfolio that is protected against adverse currency movements by purchasing currency put options will have the same return as a domestic portfolio that combines a long fixed income securities with purchased currency call options. Sec. 8.3 demonstrates how the put call parity condition for currency options can be used to create insured portfolios of foreign bonds.

Early exercise for American currency options depends on the relationship between the foreign and domestic interest rate. To see this, consider the extension of Property 10 of Sec. 7.1 to the continuous dividend case:

$$C_A[S,\tau,X] \geq C[S,\tau,X] \geq Max[0, S(t)PV[r^*,\tau] - X\,PV[r,\tau]]$$

where r^* is the continuous dividend payment, in this case paid in the form of interest on the foreign bond position. If $r^* > r$, then the European bound will be less than $S(t) - X$ and there is an incentive for the American call option to be exercised early. The extension to puts follows from observing that a currency call option giving the right to buy \$D for *\$F* at exchange rate X, is also a put option giving the right to sell \$F for *\$D* at X. When the call is in-the-money, the put is out-of-the-money. Hence, if $r > r^*$ then the European bound for the put will be less than $X - S(t)$ and there is an incentive for the American put option to be exercised early.

To see how the extension of Property 10 translates into an early exercise condition, recall that when a option is exercised early, the time value remaining in the option is given up. In practice, currency options are often exercised early, sometimes long before the stated expiration date, e.g., Bodurtha and Courtadon (1995). The incentives to exercise early follow by examining the European option price when the option is deep in the money, $S >> X$ for the call and $X >> S$ for the put. For the call, deep in-the-money implies $N[d] \rightarrow 1$ and $C \rightarrow Se^{-r^*t^*} - Xe^{-rt^*}$ which is $< S - X$ when $r^* > r$. Similarly, $N[-d] \rightarrow 1$ when $X >>$ S which gives $P \rightarrow Xe^{-rt^*} - S^{-r^*t^*}$ which is $< X - S$ when $r > r^*$. Hence, in these cases, the European option does not provide an effective lower bound for the American option. There is no arbitrage support in the market to prevent the American option from trading above the exercise boundary, $S(t) - X$.

To see the rationale for exercise in the cases without support from the European boundary, let $r^* > r$ and have $Se^{-r^*t^*} - Xe^{-rt^*} < S - X$ and the call option in the money. If the option holder borrows X at r to buy S and invests this one unit of foreign currency in the foreign bond at r^*, then on the maturity date the early exercise strategy will receive a profit of: $S(T)\, exp\{r^*t^*\} - X\, exp\{rt^*\}$. If the option holder does not exercise then the profit will be: $max[0, S(T) - X]$. Comparing these two values reveals that, if the probability of the option finishing out-of-the-money is ignored, then early exercise will be optimal. Hence, if the option is deep enough in the money that the time value has gotten close enough to zero, then it will be optimal to exercise the call option early if $r^* > r$. A similar result holds for in-the-money puts if $r > r^*$.

To compare the early exercise and hold-to-maturity values directly requires two observations. First, the call option on the exercise date is assumed to be sufficiently in-the-money that $N[d] \rightarrow 1$ and the value of the American call has been forced to the exercise boundary $S(t) - X$. In practice, there is little or no time value on deep-in-the-money options because there is no demand to buy this type of security, only those wishing to sell. Hence, this condition is often observed in practice. Second, if the option is deep-in-the-money on the exercise date, the closer this date is to the exercise date, the smaller the probability that the option will finish out-of-the-money. If the option finishes out-of-the-money then the hold-to-maturity strategy will be superior. So there is a tradeoff between the time to expiration and the difference between r^* and r required to trigger early exercise. In the jargon of Sec. 8.5, early exercise is, once again, an optimal stopping problem.

Empirical Studies of Currency Options

A number of studies have investigated the empirical validity of applying the Black-Scholes based currency option pricing model as well as competing alternatives. For example, Chesney and Scott (1989) use data on European currency options traded in Geneva to compare the Black-Scholes model with a model that allows for stochastic variance which is a mean-reverting diffusion. When an updated implied standard deviation is used to compute the Black-Scholes option price, Black-Scholes outperforms the stochastic variance model. (Black-Scholes with constant variance performed poorly). Limited evidence is found for significant mispricing. Similar results are reported by Shastri and Tandon (1986) where a number of different methods of calculating the variance are compared. Profit opportunities that were identified were found to be exhausted within one day. Using PHILX options data, Hilliard, et.al. (1991) compare the constant variance Black-Scholes model with a model that admits stochastic interest rates. While the stochastic rate model is found to be more accurate, it would have been useful if this study had reported results for Black-Scholes using implied standard deviations, as in Shastri and Tandon and other sources. Ritchken and Trevor (1999) is a recent effort along this line of research.

Other empirical studies have examined the performance of other option pricing models. For example, Bodurtha and Courtadon (1987) examine the pricing accuracy of the American option pricing model and find: "...the standard American option pricing model does not explain the pricing of foreign currency options as well as it explains the pricing of stock options. In particular...this model under-prices out-of-the money options relative to at-the-money options and in-the-money options." (p.165) As noted previously, Shastri and Tandon (1986) also report pricing shortcomings when the European model is used to price American options. Finally, along a more analytical line, Hull and White (1987) provide results for various delta, delta + gamma and delta + sigma strategies for hedging risks associated with writing foreign currency options. With reference to the discussion in Sec. 9.1, delta + gamma hedging is found to perform relatively better only when the traded option has a relatively constant implied standard deviation and a short time to expiration. Typically, delta + sigma hedging is found to produce superior results.

Questions

1. In Sec. 7.1, a verbal description was provided to describe the geometry of the evolution of the Wiener process. Construct the appropriate geometry.

2. The Black-Scholes solution to the fundamental PDE for a European call on a non-dividend paying stock requires a change of variables to transform the PDE into the general form of a parabolic PDE. Provide the steps required to specify this change of variable.

3. Extend the discussion of Sec. 8.4 to derive the solution for the perpetual put option. This requires a different specification of θ associated with the second root of the fundamental quadratic equation, e.g., Dixit and Pindyck (1994, p.143).

4. "A call option benefits from increases in the stock price and these increases can be very large. A put option benefits from stock price declines, but the stock price can only fall to zero. Therefore, if we have a put and a call on the same stock with the same terms, the put must sell for less than the call." Do you agree or disagree? Explain.

5. Given the following information:

$S = \$47 \quad X = \$45 \quad i = .12 \quad \sigma = .40$ *(annual)*

Calculate the three month call option price that is consistent with the Black-Scholes pricing model.

6. Outline the continuous time derivation of the Black-Scholes Model. What assumptions are being to derive the results? What are the limitations of applying the model to actual options prices? Under what conditions will American calls be exercised early? What early exercise conditions apply to puts?

7. Using the two portfolio approach, develop the trading strategy underlying Property 12 of chapter 7. How does the argument work when puts are used?

8. Sec. 8.4 states the following extension of Property 10 to the continuous dividend case: $C_A[S,\tau,X] \geq C[S,\tau,X] \geq Max[0, S(t)PV[r^*,\tau] - X\ PV[r,\tau]]$. Using the two portfolio approach, verify this condition.

6. Extend Property 15 of chapter 8 to demonstrate that it will not be rational to exercise an American put just prior the stock's ex-dividend date.

7. For currency options, there is a smile relationship between volatility and moneyness. Yet, implied volatilities for equity options exhibit a volatility skew, that is downward sloping in moneyness. Discuss the possible reasons for this differing behavior in implied volatility for these two commodities.

NOTES

1. The scope of the material covered in this chapter is volumous. In addition to numerous books, journals, and other academic efforts, courses in stochastic processes are part of the typical requirements for advanced degrees in mathematical statistics. Given this, it is not possible to be either fully comprehensive or adequately technical in treating this material. The development given here is aimed at introducing concepts and notions required to develop options pricing theory, no more.

2 Written for the centenary of the *Theorie de la Speculation*, Courtault et al. (2000) provides an excellent overview of Bachelier's life.

3. One of Einstein's important contributions in this area was the observation that this random behavior could be explained by the perpetual collision of the molecules between the suspending medium and the physical objects.

4. This point is debatable. The beginnings could arguably be traced to the development of the Markov chain in 1907 by A. Markov. However, this result was aimed at the solution of a specific problem. The work of Wiener involves a complete development of the mathematical foundation for the theory of Brownian motion. There is also some disagreement about when Wiener's development of a rigorous theory of Brownian motion was developed, with Cox and Miller (p.203) using 1923 as the appropriate date.

5. Malliaris and Brock (1982) survey the range of applications in economics and finance.

6. To see this is not difficult. By construction, if y is lognormally distributed then for $ln[y] = x$, x is normally distributed. A normally distributed variable is a real variable that can take values ranging over the real line from positive to negative infinity. Observing $exp\{ln[y]\} = y = exp\{x\}$, if x takes the value of minus infinity, the lowest possible value for x on the real line, then y will take a value of zero. Hence, a lognormal variable is defined on the positive half-line, ranging from 0 to positive infinity.

7. The empirical literature on the distribution of security and derivative prices is voluminous. A useful introduction is included in Duffie (1989).

8. In the form given, the OU process is a form of arithmetic Gaussian process. Unlike the geometric case, this type of process admits the possibility of negative values for X. In certain cases, this difficulty can be rationalized away by arguing that only short term options are of interest. In other words, the probability of observing negative values increases with the length of the permissible time paths. If these paths are constrained to be short, then there will only be a negligible probability of observing negative values.

9. Analytically, the admissibility of a given diffusion for arbitrage free financial pricing depends on satisfaction of conditions required for Girsanov's theorem to hold. A discussion and application of this point to using Brownian bridge process to model bond prices is provided in Cheng (1991).

10. Extending analysis to this type of process can be motivated by observing that a diffusion process can be created from the countable combination of Poisson processes.

11. Included in these conditions are the Lipschutz condition:

$$|\alpha[t,x] - \alpha[t,y]| + |\sigma[t,x] - \sigma[t,y]| \leq K\,|x - y| \qquad \textit{for some } K > 0$$

(for admissible x and y) and the restriction on growth condition:

$$|\alpha[t,x]|^2 + |\sigma[t,x]|^2 \leq K^2\{1 + |x|^2\}$$

12. Various technical conditions associated with the lemma are suppressed, e.g., restrictions on $f[\cdot]$ and $\sigma[\cdot]$. These details can be found in various sources, e.g., Arnold (1974). It should also be recognized that other approaches to differentiation and integration of stochastic functions is possible, e.g., Meyer (1976).

13. Cox (1987, p.348-9) motivates Ito's lemma using a Taylor series expansion and ignoring terms that are of order Δt.

14. Because the nominal bond price does not have a Brownian component, no cross product terms of the form $\{dB\ dP\}$ appear. This simplification is what makes Fischer's derivation uncomplicated.

15. Black (1989) discusses the process of developing and publishing the Black-Scholes formula.

16. An interesting implication of the Black-Scholes formula is the solution for pricing the perpetual option. When t approaches infinity, $N[d_1]$ and $N[d_2]$ both go to one, with the result that the discounted value of the exercise price goes to zero. The final result is that the value of the call option equals the current stock price. This seemingly incorrect result can be explained using the discussion of the perpetual option from Sec. 8.4. When there is no dividend paid on the underlying stock, then the perpetual option will never be exercised because the optimal exercise boundary occurs at infinity. Hence, because the stock price will eventually grow well beyond the exercise price and the option will be deep-in-the-money. In this case the delta will be one and the perpetual call option price will move one-for-one with the stock price. Effectively, the perpetual call option price will behave the same as the stock price. This explanation goes some, but not all, of the way to explaining the seemingly incorrect result about the price of a perpetual option on a non-dividend paying stock.

17. With some manipulation is possible to show:

$$\frac{\ln\{\frac{S}{X}\} + (r + \frac{1}{2}\sigma^2)t^*}{\sigma\sqrt{t^*}} = \frac{\ln\{\frac{S}{X\exp\{-rt^*\}}\}}{\sigma\sqrt{t^*}} + \frac{1}{2}\sigma\sqrt{t^*}$$

This form is used in Gibson (1991) and other sources including Sec. 8.3.

18 In addition, a number of other more mundane issues associated with practical application remain. For example, in some presentations of the Black-Scholes formula, e.g., Turnbull (1987), monthly or quarterly frequencies are used for r and σ. This requires minor adjustments in the form but not the substance of the formula.

19. A useful study on the use of the implied standard deviations to predict future stock price variability is Beckers (1981). In the case of currency options, Shastri and Tandon (1986) demonstrate that implied standard deviations results in significantly more accurate Black-Scholes prices than those based on historical estimates. In particular, the historical standard deviation "yields model prices which are, on an average, lower than market prices" (p.150) Similarly, Chesney and Scott (1989) also confirm that the empirical performance of Black-Scholes model using a constant variance "performs poorly".

20. Various studies have investigated the empirical validity of Black-Scholes and its immediate variants discussed in Sec. 7.3., e.g., Macbeth and Merville (1979), Bhattacharya (1980), Bookstaber (1981), Geske and Roll (1984), Sterk (1983), Wiggins (1987) and Lauterbach and Schultz (1990). Most studies find some small sources of bias in the prices, usually associated with deep-in or deep-out-of-the-money options.

21 Reference to spot price is intended to include all underlying commodities, securities and assets upon which the option could be written. This also, implicitly, includes futures and forward prices.

22. In some cases, a volatility frown is observed, e.g., Wilmott (1998, p.290).

23. Another stylized fact associated with implied volatility is "that the differences in implied volatilities becomes less pronounced as options with greater time-to-maturity are considered" (Das and Sundaram 1997, p.2).

24. In general, Black-Scholes requires path independence of the hedge portfolio in order for a closed form solution to be possible. Hence, in order to keep any modifications "simple", this property cannot be undermined. This happens, for example, where the dividend payments are uncertain or where the options are American puts or American calls on dividend paying stocks.

25. This presentation uses the equivalent representation for d_1, i.e., except for the inclusion of the dividend d_1 is unchanged. For further discussion see the endnote associated with the specification of the Black-Scholes formula in Sec. 8.2. In practice, the discrete dividend adjusted model is easy to apply; all that is required is that the stock price be adjusted by subtracting the appropriately discounted dividend.

26. The following example is taken from Gibson (1991, Chap. 5).

27. While not realistic for many commodities to assume this full carry futures pricing model, extending to the case where the cash commodity also pays a continuous carry return is straight forward. The more complicated case where the carry return is stochastic is examined in Gibson and Schwartz (1991).

28. This formulation requires that b is, at most, a function of time, but not a function of *S* or any other stochastic variable.

29. This is not technically because of margin, transactions and other costs. In addition, over time variation margin will undermine this assumption. However, for present purposes, this assumption is not severe.

30. Early exercise can also be related to exercise price adjustment. Conditions for adjustment of the exercise price, such as stock splits or mergers, are specified in the option contract. Exercise price adjustments are not considered here.

31. All proofs given in the Poitras (1998).

32. The derivation of the transition probability density associated with the absorbed process is greatly facilitated thanks to the reflection principle, e.g., Karlin and Taylor (vol.1).

33. As, the OU suffers from the defect that there is a positive probability of negative interest rates. While for short time paths such events will, almost surely, only when the process starts close to zero, this does raise difficulties for long time paths. Hence, valuation of long term options may be problematic with this type of model.

34. The foreign bond position can be illustrated by observing that $PV^*[r^*,\tau]$ is denominated in foreign currency, say *F$*. The maturity value of the bond is *F$1*. The domestic currency (*D$*) cost of acquiring this bond when the portfolio is created at time t is $PV^*\ S(t)$, where $S(t)$ is measured in domestic direct terms, *D$/F$*. The domestic currency value of the foreign bond at maturity will be $PV^*\ S(T) = S(T)$.