Financial Engineering*

Craig Pirrong
Spring, 2008

January 15, 2008

Financial Engineering

- Since the publication of the Black-Scholes-Merton model in 1973 there has been a revolution in financial markets.

- This is the "Financial Engineering" Revolution.

- Financial Engineering has been defined as "the diagnosis, analysis, design, production, pricing, and customization of solutions" to corporate financial problems.

- Most notably, financial engineering involves the creation of new derivative securities, futures, forwards, swaps, and options of various types.

The Goal of This Course

- Financial engineering is a broad and diverse subject. This course will focus on valuation issues. That is, it will focus on how to value derivative securities of all types.

- Accurate valuation requires you to understand *and use* three important, related (but distinct) disciplines.

- Stochastic calculus.

- Numerical analysis.

- Statistics and Econometrics.

Understanding the Strengths and Weaknesses of Received Analytical Techniques

- The tools pioneered by Black-Scholes-Merton are incredibly powerful, but they are not perfect.

- Some very smart people–including Scholes and Merton–have lost immense sums–billions and billions of dollars–by putting too much faith in these models.

- A good practitioner must know how the models work, their strengths, and their weaknesses.

Strengths and Weaknesses

- The strength of existing valuation methodology are its rigor and flexibility.

- The weakness is that these methods utilize mathematical tools that make assumptions about the behavior of financial prices that are inconsistent with their real world behavior.

- “When your only tool is a hammer, everything looks like a nail.”

- “The drunk looked under the streetlight for his car keys because the light was better there.”

- This course will teach you how to use a hammer, but to recognize when you’re not driving a nail.

Stochastic Calculus

- Stochastic calculus is the fundamental tool in financial engineering because the focus of our interest is on random financial prices such as stock or commodity prices.

- Stochastic calculus allows us to determine how functions of random variables behave.

- Stochastic calculus works in continuous time. It is usually easier to derive results in continuous time even though we have to discretize time in order to find solutions to the equations that result from these derivations.

Brownian Motion

- Brownian motion is the workhorse of stochastic calculus. It is the way that randomness is represented.

- Due to the nice properties of Brownian motion, it is the “hammer” that financial engineers apply to virtually every problem

- Brownian motion is a mathematical representation of a continuous time random walk.

- Some important properties of Brownian motion are (a) continuity (all sample paths are continuous); (b) the Markov property–the time τ probability distribution of $X(t)$ for $t > \tau$ depends only on $X(\tau)$ (i.e., no path dependence), (c) it is a "Martingale," [i.e., $E_\tau[X(t)] = X(\tau)$, (d) it is of quadratic variation. That is, defining $t_i = it/n$, as $n \to \infty$:

$$\sum_{j=1}^{n} [X(t_j) - X(t_{j-1})]^2 \to t$$

(e) over finite time increments t_{i-1} to t_i, $X(t_i) - X(t_{i-1})$ has mean zero and variance $t_i - t_{i-1}$.

Stochastic Integration

- Stochastic integration is different from traditional integration.

- A stochastic integral of a function $f(.)$ is defined as:

$$W(t) = \int_0^t f(\tau) dX(\tau) =$$

$$\lim_{n \to \infty} \sum_{j=1}^{n} f(t_{j-1})[X(t_j) - X(t_{j-1})]$$

  where $t_j = jt/n$.

- The key thing to note about this expression is that the function $f(.)$ must be "non-anticipatory." That is, the function in the summation is evaluated at the left hand point of the integration interval, t_{j-1}.

- Note that this integral (an "Ito integral") is defined as a limit in the mean-square-sense. That is,

$$\lim_{n\to\infty} E\{\sum_{j=1}^{n}[f(t_{j-1})(X(t_j)-X(t_{j-1})] - \int_0^t f(\tau)dX(\tau)\}^2 = 0$$

- In essence, this states that the variance of the difference between the summation term and the integral vanishes as n goes to infinity.

Ito's Lemma

- Ito's lemma is the key tool we will use to derive valuation formulae.

- Ito's lemma describes how functions of Brownian motions behave.

- The exact derivation of Ito's lemma is extremely technical. A heuristic approach sufficies for our purposes.

- Consider a function of a Brownian motion $F(.,.)$.

- Divide the time between 0 and t into N equal increments δt in length. Use Taylor's Theorem to approximate $F(.)$:

$$F(X_t,t) - F(X_0,0) = \sum_{j=1}^{N} F_t \delta t + \sum_{j=1}^{N} F_x \Delta X_{t+j}$$
$$+ \; .5 \sum_{j=1}^{N} F_{xx} (\Delta X_{t+j})^2 + \sum_{j=1}^{N} F_{tt} \delta_t^2$$
$$+ \; \sum_{j=1}^{N} F_{tx} (\delta t \Delta X_{t+j})$$

where $\Delta X_{t+j} = X(t_{t+j+1} - X_{t+j})$.

- Taking the mean-square-limit of this expression as $N \to \infty$ implies:

$$F(X_t) = F(X_0, 0) + \int_0^t F_x(X_\tau, \tau) dX(\tau) + \int_0^t [F_t + .5 F_{xx}(X_\tau, \tau)] d\tau$$

- The key trick is getting rid of the $dX(\tau)^2$ and replacing it with $d\tau$. We can do this because the mean-square-limit of X^2 is t.

More on Ito's Lemma

- By specifying a relatively general form for dX, we can rewrite Ito's lemma.

- Specifically, an "Ito Process" is:

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

- In this expression $\mu(X,t)$ is the "drift" in X and $\sigma(X,t)$ is the volatility. Moreover, dW_t is a Brownian motion.

- Ito's lemma therefore becomes:

$$F(X(t),t) = F(X(0),0) + \int_0^t [F_\tau + F_x\mu + .5F_{xx}\sigma^2]d\tau + \int_0^t F_x \sigma dW_t$$

- We will see the term $F_\tau + F_x\mu + .5F_{xx}\sigma^2$ repeatedly.

- Ito's lemma is more usually seen in a stochastic differential equation form rather than its stochastic integral equation form:

$$dF = F_x dX + F_t dt + .5F_{xx}\sigma^2 dt$$

$$dF = [F_\tau + F_x\mu + .5F_{xx}\sigma^2]dt + F_x\sigma dW_t$$

Multi-dimensional Ito's Lemma

- If we have a function of multipe stochastic variables x_i, $i = 1, \ldots, N$, there is a multi-dimensional version of the Ito equation:

$$dF = [F_t + \sum_{i=1}^{N} \mu_i F_i + .5 \sum_{i=1}^{N} \sum_{j=1}^{N} F_{ij} \sigma_{ij}] dt + \sum_{i=1}^{N} F_i \sigma_i dW_i$$

where $\sigma_{ij} = E(dW_i dW_j)$.

Contingent Claims Pricing: Arbitrage Derivation *a la* Black-Scholes

- If we make certain assumptions about the stochastic process that an underlying claim follows can use the stochastic calculus tools to determine the value of a contingent claim on this underlying (such as a forward or an option).

- In particular, if the underlying follows an Ito Process, we can show that the contingent claim's value solves a particular partial differential equation. This can be shown in two ways, both of which center on the concept of arbitrage.

- We will derive this PDE both ways. For simplicity, we will assume the underlying asset price follows a so-called "geometric Brownian Motion" (GBM):

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

where S_t is the underlying price at t, and μ and σ are constants. This is called a "stochastic differential equation" (SDE).

- In a GBM, the asset price can never become negative, and percentage changes in the asset price (returns) are normally distributed.

Key Assumptions

- We will make some key assumptions in our analysis. Specifically:

- Zero taxes.

- Zero transactions costs.

- Continuous trading. That is, markets are always open and you can trade every instant of time without impacting price.

- Constant risk free interest rate (we will loosen this assumption later.)

Deriving the Valuation PDE

- Consider any contingent claim on S. This could be a futures, a forward, an option, or something more exotic.

- Posit that the value of the contingent claim is a function $V(S_t, t)$.

- Form a portfolio consisting of Δ units of the underlying and one unit of the contingent claim. The value of this portfolio is $\Pi \equiv S_t\Delta + V(S_t, t)$.

- By Ito's lemma, the dynamics of this claim are:

$$\Delta dS_t + V_s dS_t + V_t dt + .5V_{ss} dS_t^2$$

- Substituting from the SDE for dS_t,

$$\begin{aligned} d\Pi &= (\mu S_t \Delta + \mu S_t V_s + V_t)dt \\ &+ \sigma S_t(\Delta + V_s)dW_t + .5\sigma^2 S_t^2 dt \end{aligned}$$

- Note that the last term comes from the fact that $dW_t^2 = dt$.

- We can make this portfolio riskless by setting $\Delta = -V_s$. Our equation then becomes:

$$d\Pi = V_t dt + .5\sigma^2 S_t^2 V_{ss} dt$$

- Since this portfolio is riskless it must earn the risk free interest rate. That is, $d\Pi = r\Pi dt = rS_t\Delta dt + rV(S_t, t)dt = -rS_tV_s dt + rV dt$.

- This implies:

$$-rS_tV_sdt + rV dt = V_tdt + .5\sigma^2 V_{SS} S_t^2 dt$$

- This is a second order parabolic PDE.

- This implies our valuation equation:

$$rV = V_t + rS_tV_s + .5\sigma^2 S_t^2 V_{ss}$$

- Note that this equation holds for *any* claim on S.

Boundary Conditions

- The only thing that differentiates distinct contingent claims on the same underlying (e.g., on the same stock) is their boundary conditions.

- For instance, for a call expiring at T with strike price X, we know that $V(S_T, T) = \max[S_T - X, 0]$. For a put, we know that $V(S_T, T) = \max[X - S_T, 0]$. For a forward contract, we know that $V(S_T, T) = S_T$.

- These payoff conditions are not sufficient to determine V. We need two additional boundary conditions. These are also driven by the nature of the problem. For instance, with a European call we know that $V(0, t) = 0$ and $\lim_{S \to \infty} = S - e^{-r(T-t)}X$.

- Other examples. For an American call we know that a smooth pasting condition must hold. For knock-out option (say a knock-out call) the value of the claim has to be zero at the barrier.

- Given the PDE and the barrier conditions, we can solve for the value of the claim.

A Remarkable Feature

- Note that the value of the contingent claim does *not* depend on the "true" drift of S. That is, μ does not appear in the PDE. The true drift is the expected return, and depends on the risk aversion (i.e., the preferences) of investors.

- For this reason, this is sometimes called "preference free" pricing.

- Instead of the true drift, the risk free interest rate appears in the valuation PDE.

- That is, you can value the contingent claim *as if* the expected return on the undelryying equals the risk free rate. This would be true in general if and only if all investors are risk neutral. That is why this is sometimes referred to as "risk neutral" pricing.

- This result is an artifact of hedging. Due to continuous trading and the absence of arbitrage, an investor can hedge away all risk of holding the contingent claim by trading the underlying.

- The hedge must be adjusted dynamically and continuously because V_s changes with the underlying price.

- Contingent claims are valued by *replication* (i.e., hedging) *not* by expectation. The drift affects expected payoffs, but this does not matter for valuation purposes.

- Thus, if the underlying is traded, we can pretend that we are in a risk neutral world.

- This is a wonderful feature because it is notoriously hard to estimate the true drift of asset prices.

- The alternative derivation of the pricing equation arrives at the same conclusion using a different mathematical approach.

A Convenient Change in Variables

- The foregoing PDE is somewhat cumbersome because it has non-constant coefficients (that is, the coefficients depend on S_t). Consider the following change of variables $Z_t = \ln(S_t)$.

- By Ito's lemma,

$$dZ = (\mu - .5\sigma^2)dt + \sigma dW_t$$

- Using this change in variables

$$rV = V_t + (r - .5\sigma^2)V_z + .5\sigma^2 V_{zz}$$

- Note that this equation has constant coefficients if σ is a constant. This makes numerical solution easier.

What Happens When the Underlying is Not Traded?

- The foregoing analysis depends on the assumption that the investor can hedge away the dW_t risk by trading the underlying–there is a nice linear relationship between dS_t and dW_t. What if the underlying itself is *not* traded?

- This is relevant in many contexts. For instance, if the underlying risk factor is an interest rate, the interest rate itself is not traded. As another example (that we will explore in more detail later), if volatility σ is not a constant or a function of S_t, but is instead a stochastic process, then an option price is a function of a non-traded asset–the volatility.

- Other examples: weather derivatives and power derivatives. Weather is obviously not a traded asset, but there are derivatives written on weather. Similarly, since power is not storable, you cannot create a hedge portfolio that involves holding a position in spot power–I must consume power the instant I purchase it, and cannot re-sell it even an instant later.

- We can use our hedge derivation even if the underlying isn't traded, but we will no longer be able to derive preference free results. Instead, our pricing equations will depend on the true drift. Equivalently, any pricing equation will have a "market price of risk."

Pricing with a Non-Traded Underlying

- Consider two contingent claims on some non-traded underlying x. The claim with value denoted by V expires at time T. The claim denoted by G expires at time $T' > T$.

- The SDE for x is:

$$dx_t = \phi dt + \sigma_x dZ_t$$

  where Z_t is a Brownian motion. and σ_x and ϕ are functions of x and t.

- Form a portfolio consisting of one unit of V and Δ units of G. By Ito's lemma, the dynamics of this portfolio are:

$$\begin{aligned} d\Pi &= (\Delta\phi G_x + \phi V_x + V_t + \Delta G_t + .5\Delta\sigma_x^2 G_{xx}^2 \\ &+ .5\sigma_x^2 V_{xx}^2)dt + \sigma_x(V_x + \Delta G_x)dZ_t \end{aligned}$$

- Choose Δ to make the portfolio riskless. This requires $\Delta = -V_x/G_x$.

- Since the portfolio is riskless, and requires initial investment $V - \Delta G$,

$$rV - r\Delta G = V_t + \Delta G_t + .5\Delta\sigma_x^2 G_{xx}^2 + .5\sigma_x^2 V_{xx}^2$$

- Collecting all V terms on the lhs and all G terms on the rhs:

$$\frac{rV - V_t - .5\sigma_x^2 V_{xx}}{V_x} = \frac{rG - G_t - .5\sigma_x^2 G_{xx}}{G_x}$$

- Note that we only have one equation but two unknowns (V and G). Note, however, that the lhs is a function of T (and not T') whereas the reverse is true of the rhs. This is possible if and only if both sides are independent of the maturity date. Thus, there exists some function $a(x,t)$ such that

$$\frac{rV - V_t - .5\sigma_x^2 V_{xx}}{V_x} = a(x,t)$$

- This is conventionally rewritten:

$$a(x,t) = \phi - \sigma_x \lambda(x,t)$$

- The function λ is referred to as the "market price of x risk."

- Using this definition, we derive the following PDE:

$$rV = V_t + .5\sigma_x^2 V_{xx} + (\phi - \sigma_x \lambda)V_x$$

- Note that this equation depends on the true drift of the x process–we can't employ the convenient assumption that the drift of the underlying process is equal to the risk free rate.

- Note that this derivation is more general than the earlier one because it implies the basic valuation equation for a traded asset. Note that a traded asset itself must satisfy the PDE. Thus:

$$rS = (\phi - \sigma_x \lambda)$$

- When $\phi = \mu S$ and $\sigma_x = \sigma S$, this implies:

$$\lambda = \frac{\mu - r}{\sigma}$$

- Plugging this for λ in the PDE involving $(\phi - \sigma_x \lambda)$ returns the basic valuation formula derived earlier. Note that we cannot use this trick unless x is a traded asset.

Solving PDEs Using Finite Difference Methods

- Solution of a PDE requires determination of a *function* that satisfies the relevant equation at every possible value.

- Most PDEs have no closed form solution. Even the heat equation and Black-Scholes-Merton equations require numerical approximation.

- More complicated numerical approximation schemes are required to solve PDEs with boundary conditions that are more complicated that BSM.

- Finite difference methods are the most common means to solve PDEs.

Alternative Approaches

- There are two basic finite difference schemes–explicit and implicit. The binomial model is an example of an explicit scheme.

- Explicit schemes are simpler, but (a) can face stability problems, and (b) don't converge as quickly.

- Explicit schemes are only conditionally stable. That is, for a given choice of δt, the method is stable only if δZ is sufficiently small. If you choose too big a δZ, you get numerical garbage. And I mean garbage.

- Implicit schemes are generally superior. They are unconditionally stable. For any choice of δt, you will get a non-garbage answer regardless of the coarseness of the δZ grid.

The Explicit Approach

- Although the explicit approach is deficient in many ways, it is worthwhile to discuss it to illustrate the basics of finite difference methods. It is also a flexible and easy to code approach that is useful when you need something fast.

- All finite difference schemes start with a grid. That is, the state variables and time variable are discretized.

- Consider a stock price model in which the natural log of the stock price Z is used as the state variable. Assuming constant σ and r, we know that the value of any contingent claim on this stock must satisfy the following PDE:

$$\frac{\partial V}{\partial t} + (r - .5\sigma^2)\frac{\partial V}{\partial Z} + .5\sigma^2\frac{\partial^2 V}{\partial Z^2} = rV$$

- Solution of the PDE via finite differences requires approximation of the relevant partial derivatives on the grid.

- Step 1: Create a grid. There are $I + 1$ log stock price points $i = 0, \ldots I$. Although it is not necessary, assume that the grid points are evenly spaced, with each one δZ apart. There are K time steps. Each step is δt in length. The notation is that time step $k = 0$ at expiration, and today is $k = K\delta$. This notation is used because solution involves moving backwards through time in the grid, going from values we know (at expiration) to those we don't.

- Step 2: Estimate the partial derivatives. Different schemes use different estimates.

In the explict scheme, at node $i, k+1$ of the grid:

$$\frac{\partial V_i}{\partial Z} \approx \frac{V_{i+1}^k - V_{i-1}^k}{2\delta Z}$$

where i indicates the stock price step and k indicates the time step, and V_i^k is the value of the contingent claim at node i, k.

$$\frac{\partial^2 V_i}{\partial Z^2} \approx \frac{V_{i+1}^k - 2V_i^k + V_{i-1}^k}{(\delta Z)^2}$$

$$\frac{\partial V_i}{\partial t} \approx \frac{V_i^k - V_i^{k+1}}{\delta t}$$

- Note well: at time step $k+1$ we are using values from time step k to calculate the partial derivatives.

- This has one plus and one minus. The plus is (as we will see in a bit) that the explicit method requires no solution of a system of equations.

- The minus is that we really aren't calculating a partial derivative because *we aren't holding time constant.* This is what introduces the instability in the explicit approach.

- We now plug our estimates for the partials into our PDE to get:

$$V_i^{k+1} = AV_{i-1}^k + BV_i^k + CV_{i+1}^k$$

where $A = (r - .5\sigma^2)\frac{\delta t}{2\delta Z} + .5\sigma^2\frac{\delta t}{\delta Z^2}$, $B = 1 - r\delta t - \sigma^2\frac{\delta t}{\delta Z^2}$, and $C = -(r - .5\sigma^2)\frac{\delta t}{2\delta Z} + .5\sigma^2\frac{\delta t}{\delta Z^2}$.

- We start at $k = 0$, and solve for the value of the option at time step 1, which is one time step prior to expiration. At this time step, we know the value of the option at $k = 0$ as a function of the price in the grid; that is given by the contractual features of the option. So we know everything on the right hand side of our expression, so we can solve for V_i^2 for $i = 1, \ldots I - 1$.

- For $i = 0$ and $i = I$, we need to use boundary conditions because for these values of i we need to know the value of the option at points outside our grid to solve our equation. I'll talk more about boundary conditions later.

- Since we now know the value of the option for all i at time step 1, we can proceed to time step 2. Solve the equations again, only using the values from time step 1 on the right hand side. Then go to time step 3, and so on, until we reach the valuation date.

- Couldn't be easier, eh? Just solve $I + 1$ equations (one at a time) for each time step.

The Implicit Approach

- Again start with a grid.

- In the implicit scheme, at node $i, k+1$ of the grid:

$$\frac{\partial V_i}{\partial Z} \approx \frac{V_{i+1}^{k+1} - V_{i-1}^{k+1}}{2\delta Z}$$

where i indicates the stock price step and k indicates the time step, and V_i^k is the value of the contingent claim at node i, k.

$$\frac{\partial^2 V_i}{\partial Z^2} \approx \frac{V_{i+1}^{k+1} - 2V_i^{k+1} + V_{i-1}^{k+1}}{(\delta Z)^2}$$

$$\frac{\partial V_i}{\partial t} \approx \frac{V_i^k - V_i^{k+1}}{\delta t}$$

- Using this approximation, the valuation PDE can be rewritten as:

$$AV_{i-1}^{k+1} + BV_i^{k+1} + CV_{i+1}^{k+1} = V_i^k$$

where $A = \frac{\delta t}{2\delta Z}(r - .5\sigma^2) - \frac{\delta t}{2\delta Z^2}\sigma^2$, $B = 1 + \frac{\delta t}{\delta Z^2}\sigma^2 + r\delta t$, and $C = -\frac{\delta t}{2\delta Z}(r - .5\sigma^2) - \frac{\delta t}{2\delta Z^2}\sigma^2$.

- **Note: V_{-1} and V_{I+1} are outside the grid, so we can only write these equations for $i = 1, \ldots, I-1$.**

- In matrix form, we observe:

$$\mathcal{M}_L \mathbf{v^{k+1}} = \mathbf{v^k}$$

where

$$\mathcal{M}_L =$$

$$\begin{pmatrix} A & B & C & 0 & . & . & . \\ 0 & A & B & . & . & . & . \\ . & 0 & . & . & . & 0 & . \\ . & . & . & . & B & C & 0 \\ . & . & . & 0 & A & B & C \end{pmatrix}$$

- This matrix has $I-1$ rows and $I+1$ columns.

$$\mathbf{v^{k+1}} = \begin{pmatrix} V_0^{k+1} \\ V_1^{k+1} \\ . \\ . \\ . \\ . \\ . \\ V_{I-1}^{k+1} \\ V_I^{k+1} \end{pmatrix}$$

and

$$\mathbf{v^k} = \begin{pmatrix} V_1^k \\ V_2^k \\ . \\ . \\ . \\ . \\ . \\ V_{I-2}^k \\ V_{I-1}^k \end{pmatrix}$$

Solving The Equation System

- The above shows that the solution of the PDE involves the solution of a set of linear equations.

- Unfortunately, as written, we have $I - 1$ equations (the $I - 1$ rows of the matrices) in $I + 1$ unknowns (note there are $I + 1$ columns. We need additional equations. These come from the boundary conditions. These give us the value of V at the top and the bottom of the grid.

- Boundary conditions are defined by the problem. For a European put struck at X, for instance, we know that when $S = 0$, the put value is $Xe^{-r(T-t)}$. Thus, $V_0^{k+1} = Xe^{-r(k+1)\delta t}$. Similarly, when S is very large, the value of the put is zero. Thus, $V_I^k = 0$.

- We use the boundary conditions to add two additional rows to our matrix and vector– one row corresponding to the upper boundary condition and another corresponding to the lower boundary conditon.

- Denote the new M_L matrix that includes the additonal rows as $\widehat{M}_L$:

$$\widehat{M}_L =$$

$$\begin{pmatrix} A_0 & B_0 & C_0 & D_0 & . & . & \\ A & B & C & 0 & . & . & . \\ 0 & A & B & . & . & . & . \\ . & 0 & . & . & . & 0 & . \\ . & . & . & . & B & C & 0 \\ . & . & . & 0 & A & B & C \\ . & . & 0 & A_I & B_I & C_I & D_I \end{pmatrix}$$

- The A_0, B_0, C_0, D_0, A_I, B_I, C_I, D_I come from the boundary conditions.

- Finally, we need to add two rows to the vector on the right-hand-side. This also reflects the boundary conditions. Call this new RHS vector $\hat{\mathbf{v}}^{\mathbf{k}}$.

- Our equation system now becomes:

$$\widehat{\mathcal{M}}_L \mathbf{v}^{\mathbf{k+1}} = \hat{\mathbf{v}}^{\mathbf{k}}$$

- We'll see how to adjust the matrices to reflect the boundary conditions.

Dirichlet Boundary Conditions

- Dirichlet boundary conditions set the value of the claim equal to some known number at the top and bottom of the grid.

- For instance, the value of a European put is 0 as the stock or futures price approaches infinity. The value of the Euro put is $Xe^{-r(T-t)}$ when S or F equals 0.

- Therefore, when using a Dirichlet condition for the lower boundary, set $A_0 = 1$, $B_0 = 0$, $C_0 = 0$, $D_0 = 0$, and the first row of $\hat{\mathbf{v}}^{\mathbf{k}}$ equal to the value of the claim when the underlying price is its lowest value on the grid. For instance, for a European put, set this value equal to $Xe^{-r(T-t)}$ if the lowest value of the underlying in your grid is zero.

If the lowest value in your grid is not zero (which will be the case if you use the log transform, for instance) set this value equal to $Xe^{-r(T-t)} - S_0$ where S_0 is the lowest value of the underlying price in your grid.

- Similarly, for a Dirichlet upper boundary condition, set $A_I = 0$, $B_I = 0$, $C_I = 0$, $D_0 = 1$, and the last row of $\hat{\mathbf{v}}^{\mathbf{k}}$ equal to the value of the claim when the underlying price is its lowest value on the grid. For instance, for a European put, set this value to zero.

Von Neumann Conditions

- Von Neumann conditions fix the shape of the price function at the boundaries. For example, von Neumann conditions can set the slope of the function at the boundary equal to 1, or to zero. Alternatively, they many set the second derivative at the boundary to some value, such as zero.

- As an example, for a European call, the slope of the value function should approach 1 as the stock price approaches infinity, and the slope should approach zero as the stock price approaches zero.

- **Warning: To ease the notation what follows assumes you are NOT using the log transform. I'll show you how to implement this in the log transform later.**

- Consider the upper boundary condition. We have:

$$\frac{\partial V}{\partial S} = 1$$

- Our first order approximation of this, using a one sided estimate of the derivative is:

$$\frac{V_I^{k+1} - V_{I-1}^{k+1}}{\delta S} = 1$$

- So, our upper boundary condition is:

$$V_I^{k+1} = \delta S + V_{I-1}^{k+1}$$

or

$$-V_{I-1}^{k+1} + V_I^{k+1} = -\delta S$$

- We can implement this in our matrix by setting $A_I = 0$, $B_I = 0$, $C_I = -1$, $D_I = 1$, and the last row of $\widehat{\mathbf{v^k}}$ equal to $-\delta S$.

- Our lower boundary condition can be expressed:

$$V_1^{k+1} - V_0^{k+1} = 0$$

- So we implement this by setting $A_0 = -1$, $B_0 = 1$, $C_0 = 0$, $D_0 = 0$, and the first row of $\mathbf{v^k}$ equal to 0.

- Many derivatives are nearly linear at the boundaries. Puts and calls are examples of this. So another common type of boundary condition is to set the second derivative equal to zero at the boundary.

- Our approximation of this is at the upper boundary is:

$$\frac{V_I^{k+1} - 2V_{I-1}^{k+1} + V_{I-2}^{k+1}}{\delta S^2} = 0$$

or

$$V_{I-2}^{k+1} - 2V_{I-1}^{k+1} + V_I^{k+1} = 0$$

- To implement this, set $A_I = 0$, $B_I = 1$, $C_I = -2$, $D_I = 1$, and set the last row of $\hat{\mathbf{v}}^{\mathbf{k}}$ equal to 0.

- Similarly, at the lower boundary, $A_0 = 1$, $B_0 = -2$, $C_0 = 1$, $D_0 = 0$, and the first row $\hat{\mathbf{v}}^{\mathbf{k}}$ equal to 0.

The Crank-Nicolson Appraoch

- A method that combines implicit and explicit methods–the Crank-Nicolson routine–has very desirable convergence and stability properties.

- There are some stability problems with C-N that lead some (like Duffie) to favor implicit schemes combined with extrapolation.

Implementing Crank-Nicolson

- Again start with our grid.

- In Crank-Nicolson, at node $i, k+1$ of the grid:

$$\frac{\partial V_i}{\partial Z} \approx .5\left[\frac{V_{i+1}^{k+1} - V_{i-1}^{k+1}}{2\delta Z} + \frac{V_{i+1}^{k} - V_{i-1}^{k}}{2\delta Z}\right]$$

  where i indicates the stock price step and k indicates the time step, and V_i^k is the value of the contingent claim at node i, k.

$$\frac{\partial^2 V_i}{\partial Z^2} \approx .5\frac{V_{i+1}^{k+1} - 2V_i^{k+1} + V_{i-1}^{k+1}}{(\delta Z)^2} + .5\frac{V_{i+1}^{k} - 2V_i^{k} + V_{i-1}^{k}}{(\delta Z)^2}$$

$$\frac{\partial V_i}{\partial t} \approx \frac{V_i^k - V_i^{k+1}}{\delta t}$$

- Note that the C-N approximations are averages of the implicit and explicit approximations for the partials.

- Using this approximation, the valuation PDE can be rewritten as:

$$-AV_{i-1}^{k+1} + (1-B)V_i^{k+1} - CV_{i+1}^{k+1} = AV_{i-1}^{k} + (1+B)V_i^{k} + CV_{i+1}^{k}$$

  where $A = -\frac{\delta t}{4\delta Z}(r - .5\sigma^2) + \frac{\delta t}{4\delta Z^2}\sigma^2$, $B = -\frac{\delta t}{2\delta Z^2}\sigma^2 - .5r\delta t$, and $C = \frac{\delta t}{4\delta Z}(r - .5\sigma^2) + \frac{\delta t}{4\delta Z^2}\sigma^2$.

Matrix Form

- In matrix form, we observe:

$$\mathcal{M}_L \mathbf{v^{k+1}} = \mathcal{M}_R \mathbf{v^k}$$

where

$$\mathcal{M}_L =$$

$$\begin{pmatrix} -A & (1-B) & -C & 0 & . & . & . \\ 0 & -A & (1-B) & . & . & . & . \\ . & 0 & . & . & . & 0 & . \\ . & . & . & . & 1-B & -C & 0 \\ . & . & . & 0 & -A & (1-B) & -C \end{pmatrix}$$

$$\mathbf{v^{k+1}} = \begin{pmatrix} V_0^{k+1} \\ V_1^{k+1} \\ . \\ . \\ . \\ . \\ . \\ V_{I-1}^{k+1} \\ V_I^{k+1} \end{pmatrix}$$

$$\mathcal{M}_R =$$

$$\begin{pmatrix} A & (1+B) & C & 0 & . & . & . \\ 0 & A & (1+B) & . & . & . & . \\ . & 0 & . & . & . & 0 & . \\ . & . & . & . & 1+B & C & 0 \\ . & . & . & 0 & A & (1+B) & C \end{pmatrix}$$

and

$$\mathbf{v^k} = \begin{pmatrix} V_0^k \\ V_1^k \\ . \\ . \\ . \\ . \\ . \\ V_{I-1}^k \\ V_I^k \end{pmatrix}$$

Solving The Equation System

- The above shows that the solution of the PDE involves the solution of a set of linear equations.

- Unfortunately, as written, we have $I-1$ equations (the $I-1$ rows of the matrices) in $I+1$ unknowns (note there are $I+1$ columns. We need additional equations. These come from the boundary conditions. These give us the value of V at the top and the bottom of the grid.

- Boundary conditions are defined by the problem. For a European put struck at X, for instance, we know that when $S=0$, the put value is $Xe^{-r(T-t)}$. Thus, $V_0^{k+1} = Xe^{-r(k+1)\delta t}$. Similarly, when S is very large, the value of the put is zero. Thus, $V_I^k = 0$.

- We use the boundary conditions to add two additional rows to our matrix and vector– one row corresponding to the upper boundary condition and another corresponding to the lower boundary conditon.

- Denote the new M_L matrix that includes the additonal rows as $\widehat{M}_L$:

$$\widehat{\mathcal{M}}_L =$$

$$\begin{pmatrix} A_0 & B_0 & C_0 & D_0 & . & . & \\ -A & (1-B) & -C & 0 & . & . & . \\ 0 & -A & (1-B) & . & . & . & . \\ . & 0 & . & . & . & 0 & . \\ . & . & . & . & 1-B & -C & 0 \\ . & . & . & 0 & -A & (1-B) & -C \\ . & . & 0 & A_I & B_I & C_I & D_I \end{pmatrix}$$

- The A_0, B_0, C_0, D_0, A_I, B_I, C_I, D_I come from the boundary conditions.

- Moreover we need to add two rows to the $\mathcal{M}_R$ matrix. Call this $\mathcal{M}_R$. The new top and bottom rows are all zeros.

- Finally, we need to add a new $I+1$ vector $\mathbf{r^k}$ to the right-hand-side. This vector takes the boundary conditions into account. It has zeros in rows $i = 2, \ldots I-1$. It may have non-zero values for $i = 0$ and $i = I$. These values depend on the boudnary conditions.

- Therefore, our new system becomes:

$$\hat{\mathcal{M}}_L \mathbf{v^{k+1}} = \hat{\mathcal{M}}_R \mathbf{v^k} + \mathbf{r^k}$$

Solving the PDE

- The essence of the solution technique should now be easy to understand.

- Start with what you know–the payoff to the option at expiration. For a put with strike price X, for instance, at expiration $V_i^0 = \max[Se^{Z_i^0} - X, 0]$, where S is the current stock price.

- Given V_i^0, $\forall i$, you know $\mathbf{v^0}$. Then you can solve the linear equation system for $\mathbf{v^1}$.

- Now you know $\mathbf{v^1}$, you can solve for $\mathbf{v^2}$. Continue this process until you get to $\mathbf{v^K}$.

Dirichlet Boundary Conditions for C-N

- For Crank-Nicolson, when using a Dirichlet condition for the lower boundary, set $A_0 = 1$, $B_0 = 0$, $C_0 = 0$, $D_0 = 0$, and the first row of $\mathbf{r^k}$ equal to the value of the claim when the underlying price is its lowest value on the grid. For instance, for a European put, set this value equal to $Xe^{-r(T-t)}$ if the lowest value of the underlying in your grid is zero. If the lowest value in your grid is not zero (which will be the case if you use the log transform, for instance) set this value equal to $Xe^{-r(T-t)} - S_0$ where S_0 is the lowest value of the underlying price in your grid.

- Similarly, for a Dirichlet upper boundary condition, set $A_I = 0$, $B_I = 0$, $C_I = 0$, $D_0 = 1$, and the last row of $\mathbf{r}^{\mathbf{k}}$ equal to the value of the claim when the underlying price is its lowest value on the grid. For instance, for a European put, set this value to zero.

Von Neumann Conditions

- **Warning: To ease the notation what follows assumes you are NOT using the log transform. I'll show you how to implement this in the log transform later.**

- Consider the upper boundary condition. We have:

$$\frac{\partial V}{\partial S} = 1$$

- Our first order approximation of this, using a one sided estimate of the derivative is:

$$\frac{V_I^{k+1} - V_{I-1}^{k+1}}{\delta S} = 1$$

- So, our upper boundary condition is:

$$V_I^{k+1} = \delta S + V_{I-1}^{k+1}$$

or

$$-V_{I-1}^{k+1} + V_{I}^{k+1} = -\delta S$$

- We can implement this in our matrix by setting $A_I = 0$, $B_I = 0$, $C_I = -1$, $D_I = 1$, and $\mathbf{r_I^k} = -\delta S$.

- Our lower boundary condition can be expressed:

$$V_1^{k+1} - V_0^{k+1} = 0$$

- So we implement this by setting $A_0 = -1$, $B_0 = 1$, $C_0 = 0$, $D_0 = 0$, and $\mathbf{r_0^k} = 0$.

- Many derivatives are nearly linear at the boundaries. Puts and calls are examples of this. So another common type of boundary condition is to set the second derivative equal to zero at the boundary.

- Our approximation of this is at the upper boundary is:

$$\frac{V_I^{k+1} - 2V_{I-1}^{k+1} + V_{I-2}^{k+1}}{\delta S^2} = 0$$

or

$$V_{I-2}^{k+1} - 2V_{I-1}^{k+1} + V_I^{k+1} = 0$$

- To implement this, set $A_I = 0$, $B_I = 1$, $C_I = -2$, $D_I = 1$, and set $\mathbf{r_I^k} = 0$.

- Similarly, at the lower boundary, $A_0 = 1$, $B_0 = -2$, $C_0 = 1$, $D_0 = 0$, and $\mathbf{r_0^k} = 0$.

Solving the linear equation system: The LU Decomposition

- We want to solve the following equation for $\mathbf{v^{k+1}}$:

$$\mathbf{M_L}\widehat{\mathbf{v}}^{\mathbf{k+1}} = \mathcal{M}_R \mathbf{v^k} + \mathbf{r^k}$$

- The brute force way to solve this equation is to invert $\mathbf{M_L}$. This is computationally expensive.

- There are other solution techniques that are much more efficient computationally.

- For *European* options, the LU decomposition is the best approach.

- You can decompose the square matrix $\hat{\mathbf{M}}_{\mathbf{L}}$into two other matrices, one of which has non-zero elements only on the diagonal and the sub-diagonal ($\mathbf{L}$) and another which has non-zero elements only on the diagonal and the superdiagonal ($\mathbf{U}$) such that $\hat{\mathbf{M}}_{\mathbf{L}} = LU$. Also, you can scale things so that $\mathbf{L}$ has ones on the diagonal.

- To apply the LU decomposition, define $\mathbf{q} = \mathcal{M}_R \mathbf{v}^{\mathbf{k}} + \mathbf{r}^{\mathbf{k}}$.

- Then $\mathbf{LUv} = \mathbf{q}$. One can exploit the sparseness of $\mathbf{L}$ and $\mathbf{U}$ to solve this equation for $\mathbf{v}$.

- This is computationally more efficient because due to the diagonality of L and U, they can be inverted using back-substitution.

This involves sequential solution of N equations (where N is the number of rows and columns in the matrix of interest), each with a single unknown.

- Matlab (and some other programs) use LU decomposition to invert matrices. Therefore, if you are using Matlab (or one of these other programs) you don't need to do the decomposition yourself. Just use $inv(\mathbf{M_L})$ or $\mathbf{M_L}^{-1}$.

- So far we have assumed that σ and r don't vary over time. In this case, you only have to calculate $[\mathbf{LU}]^{-1}$ once, and apply it at each time step.

- In more complicated problems, σ and r may depend on time and the state variable. In this case, A, B, and C will depend on k and i, and the LU decomposition and inverstion must be done at each time step.

Solving the linear equation system: The SOR Method

- The LU method is quick (especially with time and state independent coefficients), but is not readily applicable to American options. The Successive Over-Relaxation (SOR) method is somewhat slower, but can handle American option problems when combined with a projection step (to get PSOR, “Projected Successive Overrelaxation.)

- This method is a modification of the Gauss-Seidel method. Like G-S, it is an iterative method. One makes an initial guess, and then modifies that guess iteratively.

- On each iteration, the new value is the old value plus a correction. One iterates until the change between the new and old values becomes very small.

- SOR solves for $\mathbf{v^{k+1}}$ iteratively. For European options the procedure is as follows. First, one chooses an "over-relaxation parameter" ω, $1 \leq \omega \leq 2$. For $n = 1$, choose $\omega = 1$. Then define matrices $\mathbf{D}$, $\mathbf{L}$, and $\mathbf{U}$ so that $\mathbf{M_L} = \mathbf{D} + \mathbf{L} + \mathbf{U}$. Define $\mathbf{M}_\omega = \mathbf{D} + \omega\mathbf{L}$ and $\mathbf{N}_\omega = (1-\omega)\mathbf{D} - \omega\mathbf{U}$. Given an initial guess of $\mathbf{v_n^{k+1}}$, solve:

$$\mathbf{v_{n+1}^{k+1}} = \mathbf{M}_\omega{}^{-1}\mathbf{N}_\omega\mathbf{v_n^{k+1}} + \omega\mathbf{M}_\omega{}^{-1}\mathbf{q}$$

- In this expresion, the $\mathbf{q}$ is derived from $\mathcal{M}_R\mathbf{v^k} + \mathbf{r^k}$. It is in essence the target value.

- In this expression n is the iteration number.

- In practice, this is pretty easy. For row i, take the i'th row of $\mathcal{M}_R\mathbf{v^k} + \mathbf{r^k}$. Call this value q_i. It is the "target" that you are trying to hit. Take your initial guess for v_i. From that subtract the difference between this q_i and the i'th row of $\mathcal{M}_L v$ *and multiply the difference by the overrelaxation parameter* ω*.* The difference is essentially an "error"–it is the difference between the value at that iteration and your target value. So, under SOR new value equals old value minus ω times error. In the G-S method, $\omega = 1$.

- Formally:

$$v_{ik}^{n+1} = v_{ik}^{n} + \frac{\omega}{M_{ii}}(q_i - \sum_{j=0}^{i-1} M_{ij} v_{kj}^{n+1} - \sum_{j=i}^{I} M_{ij} v_{kj}^{n})$$

- Note you have to loop through each row $i = 1, \ldots, I-1$ and do this for each row.

So you are double looping; you are looping through the i's, and then iterating on each row.

- Usually the initial guess v_0^{k+1} is the option value at the previous time-step.

- Continue to iterate until you achieve convergence (within some user-specified error tolerance). Convergence means that the change in v from iteration n to iteration $n+1$ is small. Formally, calculate:

$$\sum_{i=1}^{I}(v_i^{n+1} - v_i^{n})^2$$

  and stop when this sum gets sufficiently small.

- Proceed to the next time step.

- Always keep track of the number of iterations until convergence. After finishing one time step, increase ω by a little bit (say .05). If the number of iterations required for convergence for that time step is smaller than for the previous step, increase ω a little more for the next time step. If the number of iterations increases, use the ω from the *previous* time step for the remainder of your analysis.

- For American options, at each iteration step, you can't use a matrix operation because it is necessary to take into account the possibility of early exercise.

- With an American option, at a given time step k for each stock price step i (starting with $i = 2$) calculate:

$$v_{ik}^{n+1} = v_{ik}^{n} + \frac{\omega}{M_{ii}}(q_i - \sum_{j=0}^{i-1} M_{ij} v_{kj}^{n+1} - \sum_{j=i}^{I} M_{ij} v_{kj}^{n})$$

- In this expression M_{ij} is the element in row i and column j of the $\mathbf{M_L}$ matrix and q_i is the element in row i of the $\mathbf{q}$ vector.

- Note that in this method, to solve for the option value at stock price node i you use the option values for nodes $1, 2, \ldots, i-1$. Immediately after solving for the value of v_{ik}^{n+1} using this method, compare this to the exercise value of the option. If the exercise value of the option exceeds v_{ik}^{n+1}, replace v_{ik}^{n+1} with the option exercise value for use in calcuating $v_{(i+1)k}^{n+1}$. If the exercise value is smaller, use the v_{ik}^{n+1} calcuated using the above formula.

Improving Accuracy:
Richardson Extrapolation

- The greater the number of time steps and asset steps, the more accurate your solution. The Crank-Nicolson method is accurate $O(\delta t^2, \delta Z^2)$.

- Increasing accuracy in this way is computationally expensive, because the number of calculations is proportional to $1/\delta t \delta Z$.

- Richardson extrapolation allows you to get accuracy $O(\delta t^2, \delta S^3)$.

- To implement RE, first solve the problem for a given number of asset steps (e.g., 20). Call the value of the option given this approach V_1 and the asset step δS_1. Then increase the number of asset steps (e.g., to 30). Call the option value using this grid V_2 and the asset step δS_2. The RE value of the option is:

$$V^* = \frac{\delta S_2^2 V_1 - \delta S_1^2 V_2}{\delta S_2^2 - \delta S_1^2}$$

Richardson Extrapolation and the Implicit Method

- One virtue of the C-N method is that it is second-order accurate even without extrapolation, whereas the implicit method is only first order accurate.

- However, the C-N method frequently exhibits some instabilities in valuations for at-the-money strikes. These instabilities are especially evident when graphing the Greeks–the Deltas and Gammas of the option.

- The implicit method does not exhibit these spurious oscillations. So how can we can our cake and eat it too? (Or as a famous trader I know always says–how can we have our cake and cookie too?) That is, how can we get second order accuracy and no spurious oscillations?

- It's easy. Just do the implicit method with Richardson Extrapolation *on the time step.*

- First estimate the value using N_1 time steps. Call this value V_1, Then estimate the value using $2N_1$ time steps. Call this value V_2. The extrapolated value is $V_E = 2V_2 - V_1$. V_E is second-order accurate and exhibits no spurious oscillations.

Jump Conditions

- Handling discrete cash flows (e.g., dividends) and some other conditions (e.g., periodic rather than continuous monitoring of a barrier for a barrier option) requires use of so-called “Jump Conditions.” They are called this because key variables (e.g., the stock price) jumps when something happens (e.g., a dividend is paid.

- Assume a dividend of size D will be paid at time t_d. Note that the value of the option must not change as a result of the dividend payment (everyone knows the dividend will be paid). Immediately before the dividend is paid, the option is worth $V(S, t_d^-)$. Immediately afterwards, it is worth $V(S-D, t_d^+)$. Thus, $V(S, t_d^-) = V(S - D, t_d^+)$.

- We address this problem as follows. Proceed backwards through the grid in the usual fashion. When you reach t_d, solve the value of the option in the usual way. Then implement the jump condition. At each asset step i at time step k (corresponding to t_d), define $\widehat{Z}_i = ln[exp(Z_i) - D]$.

- Next define:

$$i^* = \text{Int}[\frac{\widehat{Z}_i - Z_0}{\delta Z}]$$

$$\mu = \frac{Z_{i^*+1} - \widehat{Z}}{\delta Z}$$

$$V(Z_i, t_d^-) = \mu V(Z_{i^*}, t_d^+) + (1-\mu) V(Z_{i^*+1}, t_d^+)$$

- You have to be careful when i is small, as in this case $\widehat{i} < 0$. In this case, just use $i^* = 0$ and $\mu = 1$.

- You have to be clever about setting up your time steps. In certain cases, you can just move the dividend to the closest payment date (taking care to adjust by the time value of money involved in displacing the timing of the dividend). Alternatively, you can create a new time step corresponding exactly to the divident payment. Here you have to be careful when constructing your δt to make sure that at that new time step you are using the right δt when calcuating your A, B, and C coefficients.

- If the option is an American call, you have to take into account the possibility of early exercise. In this case, you need to utilize the SOR technique discussed earlier.

Martingale Methods

- There is an alternative (and equivalent) way to value derivatives. This involves use of Martingale Methods.

- In essence, Martingale Methods imply that any derivative can be valued by calculating its discounted expected cash flows under some probability measure.

- Calculating an expectation involves integrating over the relevant probability measure.

- Thus, integration methods (Gaussian, Monte Carlo) can be used to value derivatives.

- Some elegant mathematical theory–notably Kolmogorov's backward equation and the Feynman-Kac formula–show that the value function implied by calculating the expectation under the relevant probability measure must satisfy the same valuation PDE we derived using the Black-Scholes arbitrage approach. Thus, the two methods are different ways of skinning the same cat (apologies to PETA members).

- The fact that the methods are equivalent derives from the fact that the absence of arbitrage is a necessary (and sometimes sufficient) condition for the existence of the relevant probability measure required to calculate the expectation.

Martingales

- A martingale is a "zero drift" stochastic process. That is, W_t is a martingale if

$$E[W_T|W_t] = W_t \qquad \forall T \geq t$$

- Martingales have very desirable properties that facilitate solution of valuation problems.

Probability Measure

- A probability measure assigns probabilities to events.

- Formally, define a state space Ω that defines all the possible states of the world–all the things that *can* happen. An event is a group of states of the world.

- A σ-field allows specification of sets of events to which probabilities can be assigned. A σ-field $\mathcal{A}$ on Ω has the following properties (a) $\Omega \in \mathcal{A}$, (b) if $A_i \in \mathcal{A}$, then the complement of A_i, $A_i^C \in \mathcal{A}$, and (c)if $A_i \in \mathcal{A}, i = 1, \ldots, n$ then $\cup_{i=1}^{n} A_i \in \mathcal{A}$.

- The elements of $\mathcal{A}$ are called measurable sets. A probability measure associates to each measurable set a real number in $[0, 1]$. The probability measure $\mathcal{P}$ has several properties: (a) $\mathcal{P}(\Omega) = 1$; (b) if $A_i \cap A_j = \emptyset \quad \forall i \neq j$, then $\mathcal{P}(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mathcal{P}(A_i)$; and (c) $\mathcal{P}(\emptyset) = 0$.

- A probability measure quantifies the likelihood of events. A triplet $\{\Omega, \mathcal{A}, \mathcal{P}\}$ is called a proability space.

- In financial mathematics and engineering the probability measure is almost always the Gaussian (or normal) distribution. The Gaussian distribution is like the streetlight under which the drunk looks for his keys–we use it not necessarily because it is appropriate, but because the light is better there.

Equivalent Measures

- The concept of an equivalent measure is key in modern valuation theory.

- Consider a probability measure $\mathcal{P}$ that defines the probability that a particular random variable Z will take a particular value. An equivalent measure $\mathcal{Q}$ (a) has same null sets, but (b) a different mean. That is, if $\mathcal{P}(Z = Z^*) = 0$, then $\mathcal{Q}(Z = Z^*) = 0$, but $E_{\mathcal{P}}[Z] \neq E_{\mathcal{Q}}[Z]$.

- We care about probability measures for stochastic processes. Consider a stochastic process X_t with associated probability measure $\mathcal{P}$. To be specific, assume $dX_t = \gamma(X,t)dt + \sigma(X,t)dW_t$. dW_t is a Brownian motion under the probability measure $\mathcal{P}$.

- Define processes $u(X, t)$ and $\alpha(X, t)$ such that

$$\sigma(X, t)u(X, t) = \gamma(X, t) - \alpha(X, t)$$

Assume $E[exp\{.5 \int_0^t u(x, s)^2 ds\}] < \infty$. Define a process M_t as follows:

$$M_t = exp\{-\int_0^t u(s)dW_s - .5 \int_0^t u(s)^2 ds\}$$

- Given these preliminaries, we can state an important result that eases valuation problems. This is called Girsanov's theorem.

Girsanov's Theorem

- Girsanov's theorem states that if we start with a process X_t, we can create another process that has an arbitrary drift $\alpha(X,t)$ under an equivalent measure $\mathcal{Q}$. Specifically,

$$dX_t = \alpha(X,t)dt + \sigma(X,t)d\tilde{W}_t$$

where $d\tilde{W}_t$ is a Brownian motion under the probability measure

$$d\mathcal{Q} = M_t d\mathcal{P}$$

M_t is referred to as the Radon-Nikodym derivative.

- Equivalently,

$$d\tilde{W}_t = u(X,t) + dW_t$$

- In essence, Girsanov's theorem states that given an initial process, we can create an equivalent process with an arbitrary mean by adjusting simultaneously the probability measure. This adjusted probability measure is called an equivalent measure.

- Note that $\tilde{W}_t$ is a martingale under the alternative measure $\mathcal{Q}$, but not under the original measure $\mathcal{P}$. Similarly, W_t is a martingale under $\mathcal{P}$, but not under $\mathcal{Q}$.

- There are arbitrarily many equivalent measures. One special equivalent measure is an equivalent martingale measure. Under an equivalent martingale measure, discounted prices of securities are martingales. That is, for an asset price S, if interest rates are constant, under the equivalent martingale measure $\mathcal{Q}$,

$$E_{\mathcal{Q}}[e^{-r(T-t)}S_T|S_t] = S_t$$

If interest rates are stochastic,

$$E_{\mathcal{Q}}[e^{-\int_t^T r(s)ds} S_T | S_t] = S_t$$

The Link Between the Absence of Aribtrage and the Existence of an Equivalent Martingale Measure

- Girsanov's theorem is extremely useful because there is a link between the absence of arbitrage and the existence of an equivalent martingale measure. There are two key results.

- First, if there exists an "equivalent martingale measure" $\mathcal{Q}$ such that all discounted prices processes are martingales under $\mathcal{Q}$, then there are no arbitrage opportunities.

- Second, under certain technical conditions, the absence of arbitrage implies the existence of a unique equivalent martingale measure.

- In finite state space models, the absence of arbitrage always implies the existence of an equivalent measure. In continuous models (with an infinite number of states), additional technical conditions are required to ensure the existence of an equivalent measure.

- This is where Girsanov's theorem comes in. The theorem tells us how to create an equivalent measure so that asset prices rise at the risk free rate.

An Example

- Assume interest rates are constant. Consider a stock that follows the stochastic process

$$dS_t = \mu S_t dt + \sigma S_t dZ$$

where dZ is a Brownian motion.

- The solution to this SDE under the "true" measure is:

$$E_{\mathcal{P}}[S_T] = S_0 e^{\mu T} = \int_{-\infty}^{\infty} S_0 e^{(\mu - .5\sigma^2)T + \sigma\sqrt{T}Z} \frac{e^{-.5Z^2}}{\sqrt{2\pi}}$$

- (Prove that $S_T = S_0 e^{(\mu - .5\sigma^2)T + \sigma\sqrt{T}Z}$ satisfies the above SDE–Use Ito's Lemma.)

- Referring back to the statement of Girsanov's theorem, I can choose an arbitrary drift by adjusting the probability measure.

- I want a drift rS_t. Thus, if I choose:

$$u(S,t) = \frac{\mu - r}{\sigma}$$

and a new Brownian motion:

$$d\tilde{Z}_t = u(S,t) + dZ_t$$

I get:

$$dS_t = rS_t + \sigma S_t d\tilde{Z}_t$$

where $\tilde{Z}_t$ is a martingale under the equivalent measure $\mathcal{Q}$:

$$d\mathcal{Q} = \frac{e^{-.5\tilde{Z}^2}}{\sqrt{2\pi}}$$

Using the Martingale Approach

- The martingale approach allows us to determine the value of any contingent claim expiring at T as:

$$V = E_{\mathcal{Q}}[P_t(S_t)e^{-\int_0^T r(s)ds}]$$

  where $P_t(S_t)$ is the payoff to the claim at $t \leq T$.

- Once we have established the equivalent measure, we can value any contingent claim by calculating an expectation over this measure.

- Expectations involve calculations of integrals. Thus, implementing the equivalent martingal approach entails use of numerical integration. Numerical integration methods include Gaussian quadrature and Monte Carlo techniques.

Numeraires

- A “numeraire” is an asset used to discount cash flows.

- So far we have implicitly worked using a particular “numeraire” to discount cash flows in our valuation formulae. This is the “money market account” numeraire. This is an account that grows at the risk free rate of interest.

- Although the money market account is the most common numeraire, it is not the only one. Moreover, sometimes clever choice of a different numeraire can make valuation easier.

- The Numeraire Irrelevance Theorem tells us that we should get the same value for a contingent claim regardless of the numeraire we use.

- The theorem states that if P and Q are two numeraires, the value of any contingent claim is given by V:

$$V_t = P_t E_P[V_T/P_T] = Q_t E_Q[V_T/Q_T]$$

- In this expression, the notation E_Q (resp. E_P) means that the expectation is taken under a measure in which V_T/Q_T (resp. V_T/P_T) is a martingale.

- This expression suggests that when we change numeraires, we must change probability measures. Girsanov's theorem tells us how to do this

Determining the Drift Under a Given Numeraire

- Assume initially that we use the money market account $B_t = e^{rt}$ as the numeraire. There is another asset, S_{1t}, with $dS_{1t} = \mu_1 S_{1t} dt + \sigma_1 S_{1t} dW_1$, that we want to use as a numeraire. Consider the numeraire ratio $N_t = S_{1t}/B_t$. By Ito's lemma:

$$\frac{dN_t}{N_t} = \frac{dS_{1t}}{S_{1t}} - \frac{dB_t}{B_t} - \frac{dS_{1t}}{S_{1t}}\frac{dB_t}{B_t}$$

- Note that $dB_t/B_t = rdt$, and that the last term in the above expression is therefore of $o(dt)$. Thus,

$$\frac{dN_t}{N_t} = -rdt + \mu_1 dt + \sigma_1 dW_1$$

- Now Girsanov's theorem comes into play. A corollary to the theorem (there are many ways to express GT) implies that if dW_i is a Brownian motion under the measure implied by the money market numeraire, then $d\tilde{W}_i = dW_i - dW_i(dN_t/N_t)$ is a Brownian motion under the measure implied by the new numeraire.

- Note that $dW_i - dW_i(dN_t/N_t) = \sigma_1 \rho dt$, where $\rho = E(dW_1 dW_i)$.

- We can apply this to the numeraire asset. Recall that when the money market account is the numeraire, $dS_{1t} = rS_{1t}dt + \sigma_1 S_{1t} dW_1$. Thus, when S_{1t} is the numeraire, $dS_{1t} = rS_{1t}dt + \sigma_1^2 S_{1y}dt + \sigma_1 S_{1t} d\tilde{W}_1$.

- We can apply this to other assets. Consider asset two, such that when the money market account is the numeraire $dS_{2t} = rS_{2t}dt + \sigma_2 S_{2t}dW_2$. Thus, when S_{1t} is the numeraire, $dS_{2t} = rS_{2t}dt + \sigma_1\sigma_2\rho S_{2t}dt + \sigma_2 S_{2t}d\tilde{W}_2$.

An Example: The Quanto Forward

- A "quanto" is a contract with a payoff determined by an asset with a price expressed in currency A, but paid in currency B. For instance, consider a contract with a payoff equal to the FTSE 100 stock index paid in dollars. If the index is 4000 at expiration, the quanto holder gets 4000 dollars.

- It is easy to figure out the value of a forward on the FTSE paid in pounds. Assuming no dividends for simplicity, using a pound sterling money market account as the numeraire, $F_{t+\tau} = e^{u\tau} S_t$ where u is the sterling riskless rate and S_t is the FTSE spot price.

- To price the quanto, let's change numeraires from the sterling money market to the dollar money market. In doing so, we have to be careful about converting the sterling to dollars. Calling D_t the value of the sterling money market at t, and C_t the value of the dollar/sterling exchange rate, the dollar value of the old numeraire is D_tC_t, where $dC_t = \mu_C C_t dt + \sigma_C C_t dW_C$. Our numeraire ratio N_t is therefore B_t/D_tC_t. By Ito's lemma:

$$\frac{dN_t}{N_t} = rdt - udt - \mu_C dt - \sigma_C dW_C + \sigma_C^2 dt$$

- Under the old numeraire $dS_t = uS_t dt + \sigma_S S_t dW_S$. By GT, under the new measure:

$$dS_t = uS_t dt - \sigma_C \sigma_S \rho_{CS} dt + \sigma_S S_t d\tilde{W}_S$$

where $\tilde{W}_S$ is a Brownian motion under the new measure implied by the new numeraire.

- Call the quanto forward price k. This price sets the value of the forward contract to zero. Under the new measure, for a quanto expiring at $T > t$

$$0 = E[(S_T - k)/B_T] = e^{-r(T-t)} E[S_T - k]$$

- Given the process for S under the new numeraire,

$$E[S_T] = S_t e^{(u - \rho\sigma_C\sigma_S)(T-t)} S_t$$

- This implies that $k = S_t e^{(u - \rho\sigma_C\sigma_S)(T-t)} S_t$.

The Links Between the PDE and Martingale Approaches: Komogorov's Backward Equation and Feynman-Kac

- The martingale and arbitrage approaches to contingent claim valuation seem extremely different. In fact, though, they give the exact same answer.

- Two remarkable results prove that these approaches are equivalent. The first theorem–Komogorov's backward equation–is relevant when interest rates are constant. The second–the Feynman-Kac formula–is relevant when interest rates are stochastic.

- Kolmogorov's equations states that if one defines $e^{rT}u(x,t) = E_x[f(x,T)]$, where x is the Ito proces

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

then

$$ru = u_t + \mu(x,t)u_x + .5\sigma^2(x,t)u_{xx}$$

subject to the conditon $u(T,x) = f(x)$. Under the equivalent measure, $\mu(x,t) = r$. This implies:

$$ru = u_t + ru_x + .5\sigma^2(x,t)u_{xx}$$

- This is identical to the equation we used to derive the Black-Scholes formula.

- The Feynman-Kac formula extends this result to the case of stochastic interest rates.

Another Way of Showing the Equivalence of the Approaches

- We can also use the Girsanov theorem to illustrate the equivalence between the arbitrage portfolio and equivalent martingale approaches.

- Consider an asset with SDE $dS_t = \mu S_t dt + \sigma S_t dW_t$ where dW_t is a Brownian motion under the true probablity measure $\mathcal{P}$.

- We know that under the EMM $e^{-rt}S_t$ is a martingale. Note:

$$\begin{aligned} de^{-rt}S_t &= -re^{-rt}S_t + e^{-rt}dS_t = \\ & \quad e^{-rt}[-rS_t + \mu S_t dt + \sigma S_t dW_t] \end{aligned}$$

- The Girsanov theorem implies we can find another process $d\tilde{W}_t$ such that

$$d\tilde{W}_t = dW_t + dX_t$$

in which $\tilde{W}_t$ is a martingale under the probability measure $\mathcal{Q}$ with

$$d\mathcal{Q} = e^{\int_0^t X_u dW_u - .5 \int_0^t X_u^2 du} d\mathcal{P}$$

- Subsituting, we get:

$$\begin{aligned} de^{-rt}S_t &= e^{-rt}[-rS_t + \mu S_t dt \\ &+ \sigma S_t d\tilde{W}_t - \sigma S_t dX_t] \end{aligned}$$

- For this to be a martingale, it must be the case that the drift is zero. Hence:

$$S_t(-r + \mu + \sigma dX_t) = 0$$

  This requires:

$$dX_t = \frac{\mu - r}{\sigma} dt$$

- Now consider a contingent claim V. Its discounted value must also be a martingale under the equivalent measure. Ito's lemma implies the dynamics of the discounted value are:

$$\begin{aligned} de^{-rt}V &= e^{-rt}[-rV + V_t + V_s \mu S_t + .5\sigma^2 S_t^2 V_{ss}]dt \\ &+ e^{-rt}\sigma S_t V_s dW_t \end{aligned}$$

- Next substitute $d\tilde{W}_t = \frac{\mu - r}{\sigma} + dW_t$. This implies:

$$\begin{aligned} de^{-rt}V &= e^{-rt}[-rV + V_t + V_s r S_t + .5\sigma^2 S_t^2 V_s s]dt \\ &+ e^{-rt}\sigma S_t V_s d\tilde{W}_t \end{aligned}$$

- This must be a martingale, which implies:

$$-rV + V_t + V_s r S_t + .5\sigma^2 S_t^2 V_s s = 0$$

- Again–the Black-Scholes equation!

Valuing Contingent Claims Through Integration

- The Martingale approach implies that we can value contingent claims by taking an expectation. Taking an expectation involves integrating over a probability distribution.

- There are two basic integration techniques–explicit integration, and Monte Carlo integration.

- Which is appropriate depends on circumstances.

Explicit Integration

- Let's take the simplest case–a European call option with time τ to expiration struck at K. The payoff to this option is max$[S - K, 0]$. The Martingale approach implies that the value of this call is:

$$C(S, K, \tau) =$$
$$e^{-r\tau} \int_{-\infty}^{\infty} \max[Se^{(r-.5\sigma^2)\tau - \sigma\sqrt{\tau}Z} - K, 0]$$
$$\cdot \quad \frac{e^{-.5Z^2}}{\sqrt{2\Pi}}$$

- We note that we can solve for a critical value of Z, Z^* such that $S - K = 0$. This is:

$$Z^* = \frac{\ln \frac{S}{K} + (r - .5\sigma^2)\tau}{\sigma\sqrt{\tau}}$$

- Thus

$$C(S,K,\tau) = e^{-r\tau}\int_{-\infty}^{Z^*}[Se^{(r-.5\sigma^2)\tau-\sigma\sqrt{\tau}}Z - K] \cdot \frac{e^{-.5Z^2}dZ}{\sqrt{2\Pi}}$$

- Consider the term

$$\int_{-\infty}^{Z^*} Se^{-.5\sigma^2\tau-\sigma\sqrt{\tau}}\frac{e^{-.5Z^2}dZ}{\sqrt{2\Pi}}$$

- The exponent is: $-.5\sigma^2\tau - \sigma\sqrt{\tau} - .5Z^2$. We can use the trick of "completing the square to simplify this:

$$-.5(\sigma^2\tau + 2\sigma\sqrt{\tau} + Z^2) = -.5(Z + \sigma\sqrt{\tau})^2$$

- Define a new variable $Y = Z + \sigma\sqrt{\tau}$. Since $Z \leq Z^*$, $Y \leq Z^* + \sigma\sqrt{\tau} \equiv Y^*$ Thus,

$$\int_{-\infty}^{Z^*} Se^{-.5\sigma^2\tau-\sigma\sqrt{\tau}}\frac{e^{-.5Z^2}dZ}{\sqrt{2\Pi}} = \int_{-\infty}^{Y^*} S\frac{e^{-.5Y^2}dY}{\sqrt{2\Pi}}$$

- This is $SN(Y^*)$, where $N(.)$ is the cumulative normal distribution. It can be estimated with arbitrary accuracy using numerical techniques.

- Now consider the term

$$\int_{-\infty}^{Z^*} K \frac{e^{-.5Z^2} dZ}{\sqrt{2\Pi}}$$

- This is just $KN(Z^*)$.

- Therefore,

$$C(S, K, \tau) = SN(Y^*) - e^{-r\tau} KN(Z^*)$$

- This is the Black-Scholes-Merton formula.

Monte Carlo Integration

- Monte Carlo techniques are extremely common in financial applications.

- In a Monte Carlo approach, one simulates the behavior of the price of the underlying using large numbers of random draws, and then calculates the option value as a sample average of the payoffs of the option.

- Consider an option with 100 days to maturity. First draw 100 standard normal variables $z_1, \ldots, z_{100}$. This implies a series of 100 stock prices. The stock price on day j is

$$S_j = S_0 e^{(r-.5\sigma^2)(j/365)+(\sigma/\sqrt{365})\sum_{i=1}^{j} z_i}$$

- On day 100, the value of the option is: $\max[S_{100} - K, 0]$.

- Save the value of the option at expiration. Repeat this process many times (e.g., 10,000 times) and take the average of all the values you get. Discount the average to reflect the time value of money. The discounted average is the value of the option.

Path Dependent Options

- One advantage of Monte Carlo is that you can value so-called path-dependent options.

- A path-dependent option has a payoff that depends on the *path* the underlying price follows, not just the value of the underlying at expiration.

- An example of a path-dependent option is a "knock-out" call. This option becomes valueless if the value of the underlying reaches some level prior to expiration.

- In a Monte Carlo valuation approach, you can set the value of the option equal to zero on any path that crosses the "knock-out" barrier.

Increasing the Accuracy of MC

- There are a variety of techniques to improve the accuracy and speed of Monte Carlo.

- In the antithetic variable technique, in your first sample use z_i, in your second sample use $-z_i$.

- The control variable technique requires the existence of a related option that has an analytical solution. For instance, consider an "Asian" option. An Asian option has a payoff that equals the average of the underlying price over some time interval. There is no analytical solution for an Asian option with a payoff given by an arithmetic average of prices. There is, however, a analytical solution for the value of an Asian option based on a geometric average.

- In the control variate technique, calcualate the value of both the geometric and arithmetic Asian using MC. Call the value of the arithmetic option value given by the MC approximation as A. Call the value of the geometric value given by the MC as G. Call the analytical value of the geometric option G^*. Then estimate the value of the arithmetic Asian as $A + (G - G^*)$.

- Quasi-random sequences use a special algorithm to choose the z_i. An example is the Sobol' sequence. By insuring that there are fewer "gaps" and "clumps" in the z_i, quasi-random sequences allows you to use fewer random samples when estimating your option value. Due to this, the accuracy of the MC approximation is proportional to $1/M$ using qrs, instead of $1/\sqrt{M}$ using a basic random number generator, where M is the sample size.

Which Valuation Method is Best?

- Explicit integration is preferable for European options.

- Monte Carlo is frequently the best approach when (a) valuing options that have payoffs that depend on several underlying variables (e.g., stochastic volatility) and (b) certain path dependent options.

- Monte Carlo cannot handle American options.

- PDE techniques are best for American options. Moreover, PDE approaches are frequently flexible enough to value path-dependent options (such as Asians). This requires increasing the number of state variables, which raises computation (and programming) costs.

How Well Do These Approaches Work?

- These valuation approaches work well when its assumptions closely approximate reality.

- The key assumption is that the behavior of financial prices is well-approximated by an Ito Process.

- Unfortunately, there's a lot of evidence that financial prices don't behave like Ito Processes.

The Volatility Smile

- If volatility is constant, then all options on the same underlying should have the same "implied volatility." An implied vol is the volatility parameter that sets the option value given by the BSM model equal to the market price of the option.

- If volatility is a function of S and t, as is permissable with an Ito Process, then day after day, the function $\sigma(S, t)$ that best fits options prices shouldn't change.

- We know that implied volatility is not the same for all options. Indeed, there is something called the volatility smile, or volatility "smirk." That is, implied volatilities are a function of the strike price. The implied vol for at-the-money options is lower than the implied vol for options with lower strike prices, and is sometimes lower than the implied vol for options with higher strike prices.

Financial Orthodentia: Fixing the Smile ;-)

- On a given day, you can choose (using complex mathematical techniques) a function $\sigma(S,t)$ that fits the volatility smile.

- Derman and Kani first derived a method for implementing this approach based on binomial trees. Although this is popular, and widely used, it is numerically crude. It faces what as known as an "overfitting problem." There is an infinite number of $\sigma(S,t)$ functions that can fit a finite number of option prices exactly. Which one to choose? The DK approach always gives you one such function, but if you change the data (e.g., the options prices) by the tiniest amount, the DK approach will give you a completely different function. That is, the DK approach is not stable.

- There are other, very advanced approaches that avoid overfitting. These are called "inverse techniques."

- Even if you use inverse techniques, you may run into problems. This technique is good if and only if the underlying price process is an Ito process. If it isn't, this approach will not work.

- There is evidence that stock, bond, and commodity prices are not Ito processes.

- With an Ito process, using inverse techniques you should get the same $\sigma(S, t)$ function every day. In fact, you don't. Moreover, an Ito process implies that volatility depends only on the underlying price. In fact, there is a lot of evidence that volatility changes a lot and that these changes cannot be attributed solely to changes in the underlying price.

- Indeed, there is an immense body of evidence showing that volatility varies randomly over time, and that this variation is largely independent of movements in the underlying price.

- Random volatility can explain the smile.

- If volatility is random–that is, stochastic–then the Ito process-based approach will give incorrect option valuations.

Valuation With Stochastic Volatility

- Volatility is not a traded asset. Therefore, if we want to use volatility as a state variable then (a) we now have two state variables (the underlying and the volatility) and (b) we have to use a formula with a market price of risk.

- Let's specify a fairly general S volatility process:

$$d\sigma = \mu_\sigma + \nu dW_\sigma$$

  where the parameters μ_σ and ν are potentially functions of σ and t.

- Then, calling λ the market price of volatility risk, our valuation PDE for a contingent claim with value V becomes:

$$\begin{aligned} rV &= V_t + .5\nu^2 V_{\sigma\sigma} + (\mu_\sigma - \nu\lambda)V_\sigma \\ &+ rS_tV_S + .5S_t^2\sigma^2 V_{SS} + S_t\sigma\nu V_{S\sigma} \end{aligned}$$

Implementation

- This is a two-dimensional parobolic PDE. It can be solve using a variety of methods, including finite differences (especially useful for American options), Monte Carlo integration, or Fourier techniques.

- Solving the model also requires specification of the volatlity process parameters. There are several standard processes that are tractible (but perhaps not completely realistic) used for this purpose.

- Solving the model also requires knowledge of the current value of σ and estimation of the volatility process parameters. Since σ is not observed directly, these are hard problems. That is, σ is a "latent" process that requires some fancy statistics to estimate. One approach is to use discrete time (e.g., daily data) to estimate the parameters. This is feasible if the discrete time model (e.g., a GARCH model) has a continuous time limit.

- Another problem is that it is necessary to estimate a market price of risk.

- Initially, practitioners assumed $\lambda = 0$. Much empirical evidence shows that this is incorrect. If it were true, delta hedged options positions would earn the riskless rate. In fact, returns on such positions deviate substantially from the riskless rate. This implies that there is another risk premium affecting options prices. A likely candidate is volatility risk.

- λ is also an unobserved function.

- The theoretically purest way to address all these problems is to use historical options price data as well as underlying price data to estimate μ_σ, ν, and λ, and these parameters plus current options prices to estimate σ. This is a demanding process.

Jumps

- Eyeballing any financial time series one sees instances where prices seem to change discontinuously. That is, they “jump” or “gap.”

- Remember that Brownian motions have continuous sample paths–they do not exhibit jumps. Thus, Ito processes cannot capture the jumpiness observed in prices.

- One way to address this issue is to utilize so-called “jump-diffusion” models. These marry a Brownian motion process and a Poisson process.

- A Poisson process is one that exhibits no change with probability $1-\lambda dt$ and changes with probability λdt. That is, the poisson process q is $dq = 0$ with probability $1 - \lambda dt$ and $dq = 1$ with probability λdt. λ is referred to as the intensity of the jump process.

- a stock price model that incorporates the jump is:

$$dS = \mu S dt + \sigma S dW + (J - 1) S dq$$

- In this expression, J gives the magnitude of the jump. J can be deterministic (e.g., $J = .9$, indicating that the stock prices falls 10 percent during a jump) or J can be a random variable.

- Although the jump process can be specified quite generally to *describe* observed data accurately, jumps pose acute problems for valuation.

- Remember that the valuation methods that we have used so far rely on hedging and replication arguments. The jump component cannot be hedged using the underlying, however.

- If the magnitude of the jump is known and takes a single value (i.e, J is a constant) then we can construct a hedge portfolio consisting of *two* options and the underlying. This introduces a market price of risk.

- If the jump magnitude is *stochastic* things get even more complicated. If J takes on K values, we need a portfolio of $K+1$ options and the underlying to hedge all relevant risks. This injects K risk prices. Note that if J is a continuous random variable, and hence $K = \infty$, we have an infinite number of risk prices!

- Various models have been proposed to finesse this issue. Most implicitly or explicitly assume that the jump risk is diversifiable, or can be priced using the CAPM. Such models calculate values through integration. These approaches are further examples of looking for our keys under the lamppost.

- Jump models also pose serious estimation issues. That is, it is not a trivial task to estimate the parameters of a jump process.

- Jump models have become especially popular for pricing electricity derivatives because electricity prices are very jumpy. This points out a serious issue–descriptively accurate characterizations of price processes may not lead to reliable pricing models. Even if we know the probability and intensity of jumps, unless we know how the market prices these jumps we cannot derive reliable pricing models. Even if we can take expectations *under the true measure* (because we have characterized the statistical properties of the jumpy price series) we can't value unless we know the relevant probabilities *under the equivalent measure.*

- Jump models also pose problems in electricity because λ is time dependent–a jump is more likely in the summer than the fall, for instance. Moreover, the distribution of J is almost certainly time dependent–big jumps are more likely in the summer.

- Jump models provide a great illustration of the dilemmas of derivative pricing; descriptively accurate models seldom can be incorporated in the available valuation framework. Practitioners therefore face a tradeoff between descriptive accuracy and valuation feasibility.

Interest Rate Models

- One of the most important areas of derivatives modeling involves fixed income markets. Modeling derivatives on fixed income products requires modeling interest rates.

- There are two basic “flavors” of interest rate models: (a) spot rate models, and (b) Heath-Jarrow-Morton forward rate models.

- The spot rate models are much more tractible. We will focus on those.

Spot Rate Models

- Spot rate models characterize the dynamics of the instantaneous interest rate–the interest rate at which you can borrow or lend over the next instant.

- There is in fact no real world analogue to the instantaneous spot rate–it is merely a modeling convenience.

- The generic spot rate model is a single factor model of the type:

$$dr_t = \theta(r,t)dt + \sigma(r,t)dz$$

- Different models make different assumptions about $\theta(.,.)$ and $\sigma(.,.)$.

- The earliest model is the Vasicek model:

$$dr = a(b - r)dt + \sigma dz$$

- in the Vasicek model, $a > 0$, $b > 0$, and σ are constants. This model exhibits "mean reversion." That is, when $r > b$, the spot rate tends to fall; when $r < b$, the interest rate tends to rise. Thus, the interest rate reverts to the long range mean value of b.

- Interest rates can become negative in this model. This is a potentially serious problem (especially when dealing in low interest rate environments).

- The Cox-Ingersoll-Ross (CIR) model:

$$dr = a(b - r)dt + \sigma\sqrt{r}dz$$

- This “square root process” model rules out negative interest rates. It implies that interest rate volatility rises with interest rates.

- The Ho-Lee model:

$$dr = \theta(t) + \sigma dz$$

- The Hull-White model:

$$dr = (\theta(t) - ar)dt + \sigma dz$$

- This model adds mean reversion to Ho-Lee.

- Negative interest rates are possible in both Ho-Lee and Hull-White.

A Generalized Overview to Pricing Interest Rate Derivatives

- Given a spot rate model, we can use our standard pricing techniques to value any fixed income derivative.

- Call $V(r,t)$ the value of an interest sensitive contingent claim. Then:

  $$rV = V_t + .5\sigma(r,t)^2 V_{rr} + [\theta(r,t) - \sigma(r,t)\lambda(r,t)]V_r$$

- Note that we have to include a market price of risk because the spot rate is not a traded asset.

- The spot rate models discussed earlier are used largely because they are tractible, and allow closed-form solutions of this equation for certain instruments.

- In particular, it is possible to solve analytically for zero coupon bond prices using the spot rate models discussed above. A zero coupon bond that matures at T has the boundary condition $V(r, T) = 1$.

- For the spot rate models, the zero coupon bond price P_T is of the form

$$P_T(r,t) = A(t,T)e^{-B(t,T)r(t)}$$

- In the Vasicek model:

$$B(t,T) = \frac{1 - e^{-a(T-t)}}{a}$$

and

$$A(t,T) = exp[\frac{(B(t,T) - T + t)(a^2 b - .5\sigma^2)}{a^2} - \frac{\sigma^2 B(t,T)^2}{4a}]$$

- There is also a closed form solution for a European call or put option on a zero coupon bond in the Vasicek model.

- In the CIR model:

$$B(t,T) = \frac{2(e^{\gamma(T-t)} - 1)}{(\gamma + a)(e^{\gamma(T-t)} - 1) + 2\gamma}$$

$$A(t,T) = [\frac{2\gamma e^{.5(T-t)(a+\gamma)}}{(\gamma + a)(e^{\gamma(T-t)} - 1) + 2\gamma}]^{2ab/\sigma^2}$$

where $\gamma = \sqrt{a^2 + 2\sigma^2}$.

- In the Ho-Lee model, $B(t,T) = 1$ and

$$\ln A(t,T) = \ln \frac{P(0,T)}{P(0,t)} - (T-t)\frac{\partial P(0,t)}{\partial t} - .5\sigma^2 t(T-t)^2$$

- In Hull-White:

$$B(t,T) = \frac{1 - e^{-a(T-t)}}{a}$$

$$\begin{aligned}\ln A(t,T) &= \ln \frac{P(0,T)}{P(0,t)} - B(t,T)\frac{\partial P(0,t)}{\partial t} \\ &- \frac{1}{4a^3}\sigma^2(e^{-aT} - e^{-at})^2(e^{2at} - 1)\end{aligned}$$

Term Structures

- Each spot rate model implies a term structure of interest rates. The term structure is a curve that gives the interest rate as a function of maturity. The T-period interest rate is the yield on a zero coupon bond with maturity T.

- The Vasicek model allows an upward sloping, downward sloping, or "humped" term structure. This is somewhat limiting.

- The Ho-Lee and Hull-White models allow "exact" fitting of the term structure because of the "fudge factor" $\theta(t)$. One can choose $\theta(t)$ to fit term structures exactly.

- Note that since you are fitting the models to market prices, the $\theta(t)$ function implicitly includes a market price of risk.

- To calculate $\theta(t)$ in HL, collect zero prices for every maturity. Use these prices (or the prices from Eurodollar futures or FRAs) to calculate instantaneous forward rates. An instantaneous forward rate for time t is the rate that I can lock in today for borrowing at t for repayment at $t+dt$. Call the forward rate for maturity t $F(0,t)$. Then, in HL:

$$\theta(t) = F_t(0,t) + \sigma^2 t$$

- This is sometimes approximated as $\theta(t) = F_t(0,t)$, The partial derivative is estimated using finite differences.

- In HW, follow the same data collection procedure, exept use:

$$\theta(t) = F_t(0,t) + aF(0,t) + \frac{\sigma^2}{2a}(1 - e^{-2at})$$

- This is often approximated as

$$\theta(t) = F_t(0,t) + aF(0,t)$$

Calibration

- This process of fitting $\theta(t)$ is called "calibration."

- Calibration gives an exact fit to the term structure.

- This exact fit seems comforting, but in fact is dangerous. ALWAYS BEWARE A PERFECT FIT. Perfect fitting in fact implies "overfitting."

- The problem here is similar to the problem with Derman-Kani discussed earlier. If I fit $\theta(t)$ using one set of data, and then change the data only slightly, I'll get a completely different $\theta(t)$.

- Inverse problem techniques are more appropriate for this analysis.

How Well do Short Rate Models Work?

- Short rate models are extremely popular. They have some serious problems, though.

- For one thing, single factor models are of dubious validity. Principal components analysis suggests that there are multiple factors–at least three, perhaps as many as 10–driving interest rates. The three strongest factors appear to be "shift," "twist," and "hump." Thus, single factor models can't mimic the dynamics of real world interest rates.

- Multi-factor models seem to be a desirable alternative, but they are considerably more complicated to implement.

- Relatedly, if the models were correct we should see the same $\theta(t)$ functions day after day. In fact, we don't. This reflects the fact that the calibration papers over serious limitations in the models' ability to capture real world interest rate dynamics. In particular, the models cannot reliably capture the extreme steepness of the term structure at the very short-maturity section. The fitted $\theta(t)$ functions imply that the short end should "flatten out" but it usually doesn't.

- There is also substantial evidence of stochastic volatility in interest rates. The standard models don't capture this.