CHAPTER IV - INTEREST RATE MODELS

JOSEPH G. CONLON

1. Black's Model

A basic object of study in interest rate theory is the price of a zero coupon zero default risk bond such as a treasury bond. Let

(1.1)
$P(t,T) \; = \;$ value at time $t < T$ of the bond with face value 1 which matures at time T.

We can use bond prices to compute forward rates of interest. Thus for $t < T_0 < T_1$ let

(1.2)
$F(t,T_0,T_1) \; = \;$ (simple) interest rate given at time t for borrowing at T_0 with repayment at time T_1.

Evidently the no-arbitrage value of $F(t,T_0,T_1)$ is given in terms of bond prices by the formula

$$F(t,T_0,T_1) \; = \; \frac{P(t,T_0) - P(t,T_1)}{(T_1 - T_0)P(t,T_1)} \; . \tag{1.3}$$

More generally we can consider swap rates for money borrowed at time T_0 and repaid at time T_N with interest rate payments at times $T_1, T_2, .., T_N$ where $T_0 < T_1 < \cdots < T_N$. The swap rate at time $t < T_0$ is then $R(t)$ where

$$R(t) \; = \; \frac{P(t,T_0) - P(t,T_N)}{\sum_{j=1}^N (T_j - T_{j-1})P(t,T_j)} \; . \tag{1.4}$$

Observe that the forward rate $F(t,T_0,T_1)$ and the swap rate $R(t)$ are known today (time $t = 0$) from today's yield curve, but for $t > 0$ they are unknown and are therefore *random variables*.

In order to value interest rate derivatives we need to model these random variables. To see how we might go about this we consider one of the simplest interest rate derivatives- an interest rate cap on a loan. Suppose the loan is as described above, so money is borrowed at time T_0 and repaid at time T_N with interest rate payments at times $T_1, T_2, .., T_N$. The interest rate is set at time T_{j-1} for the period (T_{j-1}, T_j) to be $L(T_{j-1}, T_j)$ where $L(t,t')$ is the floating rate available at time t for borrowing at t with repayment at time $t' > t$. If K is the interest rate cap then the contribution to the cost of the cap for the interest period (T_{j-1}, T_j) is

$$(T_j - T_{j-1}) \max[L(T_{j-1}, T_j) - K, 0] \quad \text{at time } T_j \; . \tag{1.5}$$

Hence the value of the cap for a loan of 1 at time T_0 with repayment at T_N in terms of today's money is

$$\sum_{j=1}^N (T_j - T_{j-1}) \max[L(T_{j-1}, T_j) - K, 0] \; P(0,T_j) \; . \tag{1.6}$$

We can therefore think of the cap as being a sum of N caplets with each caplet corresponding to an interest rate period, so the cap is a sum of simpler derivatives. The value today of the caplet corresponding to (1.5) is given by the formula

(1.7)

$$\text{value of caplet for the period } T_{j-1} \to T_j \ = \ P(0,T_j)(T_j - T_{j-1}) E\left[\max\{L(T_{j-1},T_j) - K, 0\} \ \right] \ .$$

In Black's model he values a caplet by making a parallel with the problem of pricing a European *call* option on a stock. First he observes that $L(T_{j-1},T_j)$ is the limit of forward rates $L(T_{j-1},T_j) = \lim_{t \to T_{j-1}} F(t,T_{j-1},T_j)$. Then he models the random variables $F(t,T_{j-1},T_j), \ t < T_{j-1}$, as in the BS model by geometric Brownian motion. Thus he sets

(1.8)

$$F(t,T_{j-1},T_j) \ = \ F(0,T_{j-1},T_j)S(t) \ , \quad \frac{dS(t)}{S(t)} \ = \ \sigma dB(t) \ , 0 \le t < T_{j-1}, \quad S(0) = 1.$$

The parameter σ in (1.7) is the Black caplet volatility for the period (T_{j-1},T_j). In the parallel with the BS model for pricing stock options we have then

(1.9)

$$\begin{cases} \text{today's stock price} & = \ \text{today's forward rate for the period } T_{j-1} \to T_j \ , \\ \text{stock volatility} & = \ \text{Black caplet volatility for the period } T_{j-1} \to T_j \ , \\ \text{expiration date} & = \ T_{j-1} \ , \\ \text{strike price} \ = \ K \ . \end{cases}$$

The value of the caplet in (1.7) is then given by $P(0,T_j)(T_j - T_{j-1})$ times the BS price of a call option with risk free rate $r = 0$ and the other parameters as in (1.9). The BS hedging strategy now consists of taking a position in the forward rate for the time period (T_{j-1},T_j), which is the same as taking a position in bonds with maturities T_{j-1} and T_j. Note that we take $r = 0$ in the BS formula since the discount factor is already included in the multiplication by $P(0,T_j)$.

There is an analogous Black model for swaptions. Swaptions are a very common form of interest rate derivative since most mortgages have prepayment options, and these are in fact a type of put swaption. To see this consider a fixed rate loan at rate K which has a prepayment option at exactly one time T_0, with the remaining interest rate payments at $T_1, T_2, .., T_N$. The cost of this prepayment option is

$$\max\{K - R(T_0), 0\} \sum_{j=1}^{N} (T_j - T_{j-1}) P(t,T_j) \quad \text{at time } t < T_0 \ , \tag{1.10}$$

where $R(\cdot)$ is the swap rate (1.4). The formula (1.10) is saying that if the swap rate at time T_0 is greater than the fixed rate the borrower is currently paying then he does nothing. If it is less than the fixed rate then he pays off the loan and takes out another loan for the remaining period at the lower rate. The cost of this to the bank at time t is given by the formula (1.10). The value of the swaption is given as an expectation of (1.10) so

(1.11)

$$\text{value of swaption today} \ = \ E\left[\max\{K - R(T_0), 0\} \ \right] \sum_{j=1}^{N} (T_j - T_{j-1}) P(0,T_j) \ .$$

In Black's model the value of the swaption is estimated by making a parallel with the problem of pricing a European *put* option on a stock. He models the random variable $R(t),\ t < T_0$, of (1.4) as in the BS model by geometric Brownian motion. Thus he sets

$$R(t) \ = \ R(0)S(t) \ , \quad \frac{dS(t)}{S(t)} \ = \ \sigma dB(t) \ , 0 \le t < T_0, \quad S(0) = 1. \tag{1.12}$$

The parameter σ in (1.12) is the Black swaption volatility for this particular swap. In the parallel with the BS model for pricing stock options we have then

$$\begin{cases} \text{today's stock price} & = \ \text{today's swap rate for the given swap} \ , \\ \text{stock volatility} & = \ \text{Black swaption volatility for the swap} \ , \\ \text{expiration date} & = \ T_0 \ , \\ \text{strike price} \ = \ K \ . \end{cases} \tag{1.13}$$

The value of the swaption in (1.10) is then given by $\sum_{j=1}^N (T_j - T_{j-1})P(0,T_j)$ times the BS price of a put option with risk free rate $r = 0$ and the other parameters as in (1.13). The BS hedging strategy now consists of taking a position in the swap rate, which means a position in bonds with maturities $T_0, ..., T_N$.

The Black model is certainly a good first step in attempting to put a fair value on interest rate derivatives. However it has serious drawbacks due to its ad-hoc approach in which only volatilities estimated from market data are inputed into the model. Correlations between forward rates should clearly play an important role in pricing interest rate derivatives as well as volatilities. The simplest way to include correlation data is to model the short rate by the solution to an SDE with parameters which will be determined by both yield curve volatility and yield curve correlation.

2. Hull-White Model

In the HW model the short rate $r(t),\ t \ge 0$, is assumed to be the solution to the initial value problem for the *linear* SDE

$$dr(t) \ = \ [\theta(t) - ar(t)] \ dt + \sigma dB(t) \ , \quad r(0) = r_0 \ . \tag{2.1}$$

There are two parameters $a, \sigma > 0$ which we assume constant and a family $\theta(s),\ s \ge 0$. The parameters a, σ are first determined from yield curve correlation and volatility data. Then the family $\theta(s),\ s \ge 0$, is determined from *calibrating the model to today's yield curve*. The model is *mean reverting* since $a > 0$ so the short term interest rate can never get very large. It can however become negative, but if this happens with small probability then we expect that prices obtained from the model are still realistic.

The linear SDE (2.1) is exactly solvable just like the corresponding ODE (2.1) when $\sigma = 0$. For the ODE initial value problem we wish to solve

$$dr(t) \ = \ [\theta(t) - ar(t)] \ dt \ , \quad r(0) = r_0 \ . \tag{2.2}$$

We can rewrite (2.2) as

$$d[e^{at}r(t)] \ = \ e^{at}\theta(t) \ dt, \quad r(0) = r_0 \ . \tag{2.3}$$

Evidently the solution to (2.3) is

$$r(t) \ = \ e^{-at}r_0 + \int_0^t e^{-a(t-s)}\theta(s) \ ds \ . \tag{2.4}$$

The solution to (2.1) is similar, so

$$r(t) \;=\; e^{-at}r_0 + \int_0^t e^{-a(t-s)}\theta(s)\,ds + \sigma\int_0^t e^{-a(t-s)}\,dB(s)\;. \tag{2.5}$$

It follows from (2.5) that for $t > 0$ the random variable $r(t)$ is Gaussian with

$$\text{mean}[r(t)] \;=\; e^{-at}r_0 + \int_0^t e^{-a(t-s)}\theta(s)\,ds\,, \quad \text{variance}[r(t)] \;=\; \sigma^2\int_0^t e^{-2a(t-s)}\,ds\;. \tag{2.6}$$

We can evaluate the integral in the formula for the variance to obtain

$$\text{variance}[r(t)] \;=\; \frac{\sigma^2}{2a}\left[1-e^{-2at}\right], \quad \text{whence variance}[r(t)] \;\simeq\; \frac{\sigma^2}{2a} \text{ at large } t. \tag{2.7}$$

Since variance is bounded for all time the short term rate is bounded with high probability for all time, a desirable property since in practice we would not expect large interest rates (certainly beyond 10%) to play an important role in our model.

To see how correlations are involved in the model we find a formula for the *autocorrelation function* of $r(t),\ t \ge 0$, which is the correlation function for $r(t_1)$ and $r(t_2)$ with $0 \le t_1 \le t_2 < \infty$. Thus from (2.5) we have that

$$\begin{aligned} \text{cov}[r(t_1), r(t_2)] \;&=\; \sigma^2 E\left[\left\{\int_0^{t_1} e^{-a(t_1-s)}\,dB(s)\right\}\left\{\int_0^{t_2} e^{-a(t_2-s)}\,dB(s)\right\}\right] \\ &=\; \sigma^2\int_0^{t_1} ds\; e^{-a(t_1-s)-a(t_2-s)} \;=\; \frac{\sigma^2}{2a}\left[e^{-a(t_2-t_1)} - e^{-a(t_2+t_1)}\right]\;. \end{aligned} \tag{2.8}$$

From (2.7), (2.8) we conclude that at large t_1 the coefficient of correlation $\rho[r(t_1), r(t_2)]$ between $r(t_1)$ and $r(t_2)$ is given as

$$\rho[r(t_1), r(t_2)] \;\simeq\; e^{-a(t_2-t_1)}\;. \tag{2.9}$$

We see from (2.9) that the parameter a has the dimensions of inverse time and that the short rates $r(t_1), r(t_2)$ are close to independent if $t_2 - t_1 >> 1/a$, so $1/a$ is the correlation timescale for the yield curve. We can now estimate from (2.7), (2.9) what are reasonable values for the parameters a, σ of the HW model. Assuming that the standard deviation for the short rate should be around 5% then (2.7) gives us that $\sigma/\sqrt{2a} \simeq 0.05$. Assuming that short term rates are close to independent after $5 - 10$ years gives us $1/a \simeq 5$. We conclude then that reasonable values of a, σ are $a \simeq 0.2, \sigma \simeq 0.03$. Hence in HW we expect the value of a to be an order of magnitude (i.e. roughly 10 times) larger than the value of σ when we take time in years.

In order to price interest rate derivatives we need to give a formula for bond prices in terms of the short rate $r(\cdot)$. A fundamental assumption of short rate models is that bond prices at time t are a *deterministic function* of the short term rate r at time t. The bond price is given by the *discount formula*

$$P(t,T) \;=\; P(r(t),t,T) \quad \text{where } P(r,t,T) \;=\; E\left[\exp\left\{-\int_t^T r(s)\,ds\right\}\;\middle|\; r(t)=r\right]\;. \tag{2.10}$$

To calibrate the short rate model to today's yield curve we shall need to choose the parameters of the model so that

$$(2.11) \qquad P(0,T) \ = \ E\left[\exp\left\{-\int_0^T r(s)\ ds\right\} \ \Big|\ r(0)=r_0\right] \quad \text{for all } T>0,$$

where the LHS of (2.11) is today's price of a zero coupon bond with face value 1 and maturity T. The value of r_0 is today's short rate.

The calibration of the HW model to today's yield curve is fairly simple because the short rate $r(t)$ can be written as a sum

$$(2.12) \ \ r(t) \ = \ r^*(t)+\alpha(t), \quad \text{where } dr^*(t) \ = \ -ar^*(t)dt+\sigma dB(t), \quad r^*(0) \ = \ 0.$$

The function $\alpha(t),\ t\ge 0$, is deterministic and satisfies

$$(2.13) \qquad d\alpha(t) \ = \ [\theta(t)-a\alpha(t)]\ dt\ , \quad \alpha(0) \ = \ \alpha_0\ .$$

It is easy to see that if $r^*(\cdot)$ satisfies (2.12) and $\alpha(\cdot)$ satisfies (2.13) with $\alpha_0=r_0$ then $r(\cdot)$ satisfies (2.2). We can view the decomposition (2.12) of the short rate as into a part $r^*(\cdot)$ which is entirely due to yield curve volatility, and a deterministic part obtained from today's yield curve. In particular if there is *zero* yield curve volatility then today's forward rates are frozen in time. The instantaneous forward rate at time t is $\alpha(t)$ and today's bond price is given by the discount formula

$$(2.14) \qquad P(0,T) \ = \ \exp\left\{-\int_0^T \alpha(s)\ ds\right\}\ .$$

From (2.11), (2.12) we have more generally that

$$(2.15) \qquad P(0,T) \ = \ g(\sigma,a,T)\exp\left\{-\int_0^T \alpha(s)\ ds\right\}$$

where the function g of the three variables σ,a,T is given by the expectation

$$(2.16) \qquad g(\sigma,a,T) \ = \ E\left[\exp\left\{-\int_0^T r^*(s)\ ds\right\} \ \Big|\ r^*(t)=0\right]\ .$$

Since $g(0,a,T)=1$ we get the formula (2.14) in the case of zero yield curve volatility.

We wish now to develop a lattice model for the SDE (2.12) which $r^*(t),\ t\ge 0$, satisfies. Our lattice will consist of integer points (m,j) with $m=0,1,2,...,$ and $|j|\le\min[m,J]$, where J is some fixed integer which we need to choose appropriately depending on a,σ. We associate with each (m,j) a time $t=m\Delta t$ and a value $r^*=j\Delta r$. In one time step Δt a walk on the lattice can go from (m,j) to one of three points $(m+1,j),\ (m+1,j+1),\ (m+1,j-1)$ if $|j|<J$. At a boundary point (m,J) a walk can go from (m,J) to $(m+1,J),\ (m+1,J-1),\ (m+1,J-2)$ with a similar formula for $(m,-J)$. The transition probabilities are defined by

$$(2.17) \qquad \begin{cases} (m,j)\to(m+1,j+1) & \text{with probability } \ p_u(j)\ , \\ (m,j)\to(m+1,j) & \text{with probability } \ p_s(j)\ , \\ (m,j)\to(m+1,j-1) & \text{with probability } \ p_d(j) \end{cases}$$

if $|j| < J$. If $j = J$ then they are defined by

$$\begin{cases} (m,J) \to (m+1,J) & \text{with probability} \quad p_u(J) \ , \\ (m,J) \to (m+1,J-1) & \text{with probability} \quad p_s(J) \ , \\ (m,J) \to (m+1,J-2) & \text{with probability} \quad p_d(J) \ . \end{cases} \tag{2.18}$$

Note that the SDE (2.12) is *reflection invariant* so the random variable $r^*(t)$ has pdf which is symmetric about 0. We build this reflection property into the lattice model by requiring that

$$p_s(j) \ = \ p_s(-j), \quad p_u(j) \ = p_d(-j) \quad \text{for } |j| \le J. \tag{2.19}$$

We find values for the probabilities p_u, p_s, p_d by equating the zeroth, first and second moments of the increment $r^*(t+\Delta t) - r^*(t)$ in the discrete and continuous models. For the discrete model the three moments are given by

$$\begin{cases} \text{zeroth moment :} & p_u(j) + p_s(j) + p_d(j) = 1 \ , \\ \text{first moment :} & [p_u(j) - p_d(j)]\Delta r \ , \\ \text{second moment :} & [p_u(j) + p_d(j)](\Delta r)^2 \ . \end{cases} \tag{2.20}$$

To get the first and second moments in the continuous case we observe from (2.12) that

$$r^*(t+\Delta t) - r^*(t) \ \simeq \ -ar^*(t)\Delta t + \sigma\sqrt{\Delta t}\ \xi \qquad \text{where } \xi \text{ is standard normal.} \tag{2.21}$$

We have then on making the approximation (2.21) that

$$\begin{cases} E[\ r^*(t+\Delta t) - r^*(t)\ ] & = \ -ar^*(t)\Delta t \ , \\ E[\ \{r^*(t+\Delta t) - r^*(t)\}^2\ ] & = \ \sigma^2\Delta t + [ar^*(t)\Delta t]^2 \ , \\ E[\ \{r^*(t+\Delta t) - r^*(t)\}^3\ ] & = \ -3a\sigma^2 r^*(t)[\Delta t]^2 - [ar^*(t)\Delta t]^3 \ . \end{cases} \tag{2.22}$$

Equating the first and second moments in the discrete case (2.20) and in the continuous case (2.22) we have that

$$\begin{cases} [p_u(j) - p_d(j)]\Delta r & = \ -aj\Delta r\Delta t \ , \\ [p_u(j) + p_d(j)](\Delta r)^2 & = \ \sigma^2\Delta t + [aj\Delta r\Delta t]^2 \ . \end{cases} \tag{2.23}$$

Evidently (2.23) has generally a unique solution for $p_u(j), p_d(j)$ and then we can find $p_s(j)$ from the first equation of (2.20). We still need to decide on suitable values for $\Delta r, \Delta t$. Note that since $p_u(j) + p_d(j) \le 1$ the second equation of (2.23) implies that $\sigma^2\Delta t/(\Delta r)^2 \le 1$, so as in the forward Euler method we must take time discretization to be the same order as the square of the space discretization. HW choose a value consistent with this namely

$$\text{HW choice :} \quad \sigma^2\Delta t/(\Delta r)^2 \ = \ 1/3 \ . \tag{2.24}$$

This is actually a minor thing but the reason the value (2.24) is taken is so that third order discrete and continuous moments agree if (2.24) holds. To see this note from (2.22) that the two third order moments agree to leading order if

$$[p_u(j) - p_d(j)](\Delta r)^3 \ = \ -3a\sigma^2 j\Delta r[\Delta t]^2 \ . \tag{2.25}$$

Then the first equation of (2.23) implies (2.25) provided $\Delta r, \Delta t$ are related by (2.24).

We take now the HW value (2.24) for the ratio and then (2.23) becomes

$$\begin{cases} p_u(j) - p_d(j) &= -aj\Delta t \ , \\ p_u(j) + p_d(j) &= 1/3 + (aj\Delta t)^2 \ . \end{cases} \tag{2.26}$$

Combining the equations (2.26) with the first equation of (2.20) we conclude that

$$\begin{cases} p_u(j) &= 1/6 + [(aj\Delta t)^2 - aj\Delta t]/2 \ , \\ p_d(j) &= 1/6 + [(aj\Delta t)^2 + aj\Delta t]/2 \ , \\ p_s(j) &= 2/3 - (aj\Delta t)^2 \ . \end{cases} \tag{2.27}$$

The formula (2.27) holds provided $|j| < J$. The case $j = J$ is dealt with similarly. The first and second moment equations become now

$$\begin{cases} [p_s(J) + 2p_d(J)]\Delta r &= aJ\Delta r\Delta t \ , \\ [p_s(J) + 4p_d(J)](\Delta r)^2 &= \sigma^2\Delta t + [aJ\Delta r\Delta t]^2 \ . \end{cases} \tag{2.28}$$

Hence we have from (2.28) on using the HW formula (2.24) again that

$$\begin{cases} p_u(J) &= 7/6 + [(aJ\Delta t)^2 - 3aJ\Delta t]/2 \ , \\ p_d(J) &= 1/6 + [(aJ\Delta t)^2 - aJ\Delta t]/2 \ , \\ p_s(J) &= -1/3 - (aJ\Delta t)^2 + 2aJ\Delta t \ . \end{cases} \tag{2.29}$$

In order to ensure that all values for the probabilities in (2.27), (2.29) are non-negative we shall need to limit the size of J. From (2.27) we see that we need to have $aJ\Delta t < \sqrt{2/3}$. We can write $p_s(J)$ in (2.29) by completing the square as

$$p_s(J) \ = \ 2/3 - [aJ\Delta t - 1]^2 \quad \text{whence } |aJ\Delta t - 1| < \sqrt{2/3} \ . \tag{2.30}$$

Observe that the condition from (2.30) forces J to be very large as well as being bounded above so we have in all that

$$aJ\Delta t < \sqrt{2/3}, \quad 1 - \sqrt{2/3} \ \le aJ\Delta t \ \le \ 1 + \sqrt{2/3} \ . \tag{2.31}$$

We can take the

$$\text{HW choice}: \quad J \ = \ J_{\text{HW}} \ = \ [1 - \sqrt{2/3}]/a\Delta t \simeq .184/a\Delta t \ , \tag{2.32}$$

in which case all the probability values (2.27), (2.29) are positive.

Next we wish to argue that we can actually take a much smaller value of J than is given in (2.32) without significant loss of accuracy in computing interest rate derivative values. We use the "3 standard deviations rule" that we used in Chapter I to reduce the infinite interval $-\infty < x < \infty$ for the parabolic PDE which we needed to solve for pricing stock options to a finite interval $a < x < b$. From (2.7) we see that the standard deviation of $r^*(t)$ is less than $\sigma/\sqrt{2a}$ no matter how large t is and the mean of $r^*(t)$ is zero. Hence $r^*(t)$ takes values larger than $3\sigma/\sqrt{2a}$ in absolute value with very small probability-less than .1%. Thus a good interval for the range of r^* is $-3\sigma/\sqrt{2a} < r < 3\sigma/\sqrt{2a}$, whence we should set $J\Delta r = 3\sigma/\sqrt{2a}$. Using again (2.24) we conclude that the alternative value

$$J \ = \ J_{\text{ALT}} \ = \ \left\{ \frac{3}{2a\Delta t} \right\}^{1/2} \tag{2.33}$$

should give the same accuracy as the HW value (2.32). Note that $J_{\text{ALT}} << J_{\text{HW}}$ if $a\Delta t$ is small. Of course we cannot use the boundary probability values (2.29) when we take $J = J_{\text{ALT}}$ since we have seen that $p_s(J) < 0$ in that case. Our intuition helps us here also since we expect that the random walk on the HW tree corresponding to a solution of the SDE (2.12) is very unlikely to reach the values $r^* = \pm J\Delta r$. Hence the actual boundary condition we set really makes little difference in pricing the values of derivatives so we can choose anything reasonable. One natural choice is to use *reflecting boundary conditions*,

$$\begin{cases} p_u(J) &= 0 , \\ p_d(J) &= 0 , \\ p_s(J) &= 1 . \end{cases} \tag{2.34}$$

Before we can price interest rate derivatives we need to calibrate the HW model to today's yield curve, which means setting the values for the function $\alpha(t),\ t \ge 0$, from the equation (2.11) or (2.15). We set

$$P^m \ = \ P(0, m\Delta t), \quad m = 1,2,..., \qquad \alpha^m \ = \alpha(m\Delta t), \quad m = 0,1,2,... \tag{2.35}$$

In (2.35) we assume we know the bond prices $P^m, m = 1,2,.,$ but in actual practice bond prices for every maturity date $T = m\Delta t$ are not available. We shall therefore need to use some *interpolation* procedure to obtain the values of P^m for all $m = 1,2,..,$ from the quoted values available for some m. We can easily find α^0 from (2.11) as

$$P^1 \ = \ \exp[-\alpha^0\Delta t] \quad \text{implies } \alpha^0 = -\log[P^1]/\Delta t > 0 \ \text{ since } P^1 < 1 . \tag{2.36}$$

To calculate α^1 we approximate the integral of the short term rate by a Riemann sum,

$$\int_0^{2\Delta t} r(t)\ dt \ \simeq \ \alpha^0\Delta t + [\alpha^1 + r^*(\Delta t)]\Delta t . \tag{2.37}$$

Then we use the fact that $r^*(\Delta t) = \Delta r$ with probability $p_u(0) = 1/6$ and $r^*(\Delta t) = 0$ with probability $p_s(0) = 2/3$. Finally we have $r^*(\Delta t) = -\Delta r$ with probability $p_d(0) = 1/6$. We conclude then from (2.37) that

$$P^2 \ = \ E\left[\exp\left\{-\int_0^{2\Delta t} r(t)\ dt\right\}\right] \ = \ e^{-(\alpha^0+\alpha^1)\Delta t}\left[\frac{2}{3} + \frac{1}{6}e^{\Delta r\Delta t} + \frac{1}{6}e^{-\Delta r\Delta t}\right] . \tag{2.38}$$

Since we have already computed α^0 we can compute the value of α^1 explicitly from (2.38).

We wish now to systematize our method of computing the α^m from bond prices. To carry this out we introduce the notion of an Arrow-Debreu (AD) security. On an historical note both Arrow and Debreu won the Nobel prize in economics. For each lattice site (m, j) we define the AD security Q^m_j by

$$Q^m_j \ = \ E\left[\exp\left\{-\int_0^{m\Delta t} r(t)\ dt\right\} \ ; \ r^*(m\Delta t) = j\Delta r\right] . \tag{2.39}$$

Thus we are restricting ourselves in (2.39) to paths $[m', X(m')],\ m' = 0,..,m$, on the lattice which have endpoint $X(m) = j$. Then Q^m_j is the sum over all such paths of the exponential times the probability of the path. In financial terms we can think of Q^m_j as the value of a *virtual security* which pays 1 if the short term rate at time $m\Delta t$ is $\alpha^m + j\Delta r$, and otherwise 0. Note that the expectation in (2.39) is *not* a conditional expectation. We can write the bond price at time $(m+1)\Delta t$ in terms of the Q_j^m as

$$P^{m+1} \ = \ \sum_{j=-\min\{m,J\}}^{\min\{m,J\}} Q_j^m \exp[-(\alpha^m + j\Delta r)\Delta t] \ . \tag{2.40}$$

Evidently we can factor $\exp[-\alpha^m \Delta t]$ out of the RHS of (2.40) and what is left depends only on $\alpha^0, ..., \alpha^{m-1}$. Hence if we have already computed these and also the $Q_j^m, \ |j| \le \min\{m, J\}$, we can compute α^m explicitly from (2.40).

The AD securities satisfy a recurrence relation corresponding to writing

$$\int_0^{(m+1)\Delta t} r(t) \ dt \ = \ \int_0^{m\Delta t} r(t) \ dt + \int_{m\Delta t}^{(m+1)\Delta t} r(t) \ dt \ . \tag{2.41}$$

This yields the recurrence

(2.42)

$$Q_j^{m+1} \ = \ Q_j^m \exp[-(\alpha^m + j\Delta r)\Delta t] p_s(j) + Q_{j+1}^m \exp[-(\alpha^m + (j+1)\Delta r)\Delta t] p_d(j+1) + Q_{j-1}^m \exp[-(\alpha^m + (j-1)\Delta r)\Delta t] p_u(j-1) \ ,$$

provided (m, j) is not close to the boundary of the HW lattice. The three terms on the RHS of (2.42) correspond to the three possible positions of the random path on the lattice at time m which terminates at time $m+1$ on the site $(m+1, j)$. The three positions are clearly $(m, j), (m, j+1), (m, j-1)$. Close to the boundary we get different recurrence formulas. In the HW version (2.29) of boundary probabilities the recurrence formula for Q_j^{m+1} with $j = J-2$ is a sum of four terms corresponding to the points $(m, J), (m, J-1), (m, J-2), (m, J-3)$ since there is a finite transition probability of going from any of these points to $(m, J-2)$ in one time step. Thus $(m, J) \to (m+1, J-2)$ with probability $p_d(J)$ and $(m, J-3) \to (m, J-2)$ with probability $p_u(J-3)$. Hence in the recurrence algorithm for AD securities we need to include the various boundary cases as well as the general formula (2.42) which is valid sufficiently far from the boundary. Note that we set $Q_0^0 = 1$ to begin the computation of the AD securities from the recurrence formulas.

Once we have the recurrence relations for the AD securities we can combine these with the bond price equation (2.40) to find the values of all the α^m. Thus if we have already computed $\alpha^0, ..., \alpha^{m-1}$ and the corresponding $Q_j^m, \ |j| \le \min\{m, J\}$, then we can compute α^m explicitly from (2.40). This enables us to compute the Q_j^{m+1} from the recurrence relations (2.42) etc and again using (2.40) with m replaced by $m+1$ we compute α^{m+1} and so on. We have therefore given an algorithm for calibrating the HW model to today's yield curve.

We have already pointed out the need for interpolation to obtain values for all $P^m, \ m = 0, 1, ...,$ from bond prices observed in the market. Let us consider what a reasonable value for Δt is. Initially we are inclined to take $\Delta t = .25$ corresponding to the three month treasury bill. However we really need to take Δt smaller than this. To see why let us take $\sigma \simeq .015$ in which case (2.24) implies that $\Delta r = \sigma\sqrt{3\Delta t} \simeq .013$. Hence for this lattice model the smallest jump in interest rate is more than 1%, which is large so we should take Δt smaller than .25 to reduce the size of the jump. We can interpolate in various ways, so we could for example interpolate observed bond prices or interpolate yields on observable bond prices. Another possibility is to interpolate observable forward rates. We have also discussed in Chapter I two kinds of interpolation -linear and spline interpolation. In our context we shall find that the graph of the function $\alpha(\cdot)$ is continuous when we use spline interpolation but has lots of discontinuities when we use linear interpolation. We can see why that is the case from the formula (2.15). This formula shows that the function $\alpha(\cdot)$ depends on the derivative of the function $T \to P(0,T),\ T > 0$. If we use linear interpolation to obtain this function then its derivative will have discontinuities, which in turn gives rise to discontinuities in the function $\alpha(\cdot)$. Because spline interpolation gives a continuously differentiable function $T \to P(0,T),\ T > 0$, the corresponding function $\alpha(\cdot)$ will be continuous.

Finally we point out some advantages and some disadvantages of the HW model.

Advantage: Mean reversion-short term rate gets large only with very small probability.

Advantage: Calibration to zero yield curve explicit (no need for equation solvers).

Disadvantage: Short term rate can become negative (but with small probability).

Disadvantage: Not clear how to compute correlation and volatility parameters a, σ from market volatility data.

3. Valuation of interest rate derivatives

We already mentioned that in short rate models bond prices are deterministic functions of (r,t,T) as given by the discount formula (2.10). The function $P(r,t,T)$ on the HW lattice becomes a function of lattice sites. Thus suppose $M\Delta t = T$ for some integer T and $m\Delta t = t$ for some integer $m \le M$. Then if $r = \alpha^m + j\Delta r$ we have $P(r,t,T) \simeq P(j,m,M)$ where the function $P(j,m,M)$ is to be computed using a recurrence backward in time. The fundamental recurrence relation is given by the discount formula

$$V(m-1,j) \ = \ [p_u(j)V(m,j+1)+p_s(j)V(m,j)+p_d(j)V(m,j-1)]\exp[-r_j^{m-1}\Delta t]\ , \tag{3.1}$$

where r_j^{m-1} is the short term rate associated with the lattice site $(m-1,j)$ so $r_j^{m-1} = \alpha^{m-1}+j\Delta r$. To compute the values of $P(j,m,M)$ we solve (3.1) backwards in time for $m \le M$ with the terminal condition $V(M,j) = 1$ for all j. Then we have that $P(j,m,M) = V(m,j)$ for $0 \le m \le M$. Observe that we can check if our calibration to today's yield curve is correct by testing if our bond values computed from (3.1) satisfy $P(0,0,M) = P^M$, where P^M is the current value of the bond with maturity T.

We have observed that the value of an interest rate cap is a sum of caplets, each corresponding to an interest rate period. In Black's model we computed the value of each caplet (1.7) and summed them up. Here we can do everything together by using the backward recurrence

$$V(m-1,j) \ = \ [\{r_j^{m-1}-K\}^+\Delta t+p_u(j)V(m,j+1)+p_s(j)V(m,j)+p_d(j)V(m,j-1)]\exp[-r_j^{m-1}\Delta t]\ , \tag{3.2}$$

with the terminal condition $V(M,j) = 0$ for all j. The value of the interest rate cap is then $V(0,0)$. Note that the caplets corresponding to (3.2) are for a time period Δt. If interest is paid only after several time periods Δt then we get a rather more complicated formula.

The value of the put swaption described in §1 can be carried out similarly. Letting $T_n = M_n\Delta t,\ n = 0,..,N$, for integers $M_0 < M_1 < \cdots < M_N$, then the swap rate $R(T_0)$ defined by (1.4) is given by

$$\text{(3.3)}\quad \text{swap}(M_0, j) \ = \ \frac{P(j, M_0, M_0) - P(j, M_0, M_N)}{\sum_{n=1}^{N}(T_n - T_{n-1})P(j, M_0, M_j)} \quad \text{where } P(j, M_0, M_0) = 1 \ .$$

From (1.10) the value of the swaption at time T_0 is given by the formula
(3.4)

$$V(M_0, j) \ = \ \max\{K - \text{swap}(M_0, j), 0\}\sum_{n=1}^{N}(T_n - T_{n-1})P(j, M_0, M_n) \quad \text{for all } j.$$

Now the value of the swaption is obtained by using the discount recurrence (3.1) for $0 \le m \le M_0$ with terminal date given by (3.4). The value of the swaption today is then $V(0, 0)$. This is the same as the expectation

(3.5) value of swaption today =

$$E\left[\ \max\{K - \text{swap}(M_0, \cdot), 0\}\ \exp\left\{-\int_0^{T_0} r(t)\ dt\right\}\sum_{n=1}^{N}(T_n - T_{n-1})P(\cdot, M_0, M_n)\ \right] .$$

Note that the value of the swaption today is *not* the same as

(3.6)

$$E\left[\ \max\{K - \text{swap}(M_0, \cdot), 0\}\ \right]\ E\left[\ \exp\left\{-\int_0^{T_0} r(t)\ dt\right\}\sum_{n=1}^{N}(T_n - T_{n-1})P(\cdot, M_0, M_n)\ \right]$$

$$= \ E\left[\ \max\{K - \text{swap}(M_0, \cdot), 0\}\ \right]\ \sum_{n=1}^{N}(T_n - T_{n-1})P(0, T_n) \ .$$

The formula (3.6) is analogous to the formula (1.11) of Black's model for pricing swaptions.

We can also use Monte-Carlo methods to value interest rate derivatives. To do this we need to generate random walks through the HW lattice. A walk then is a sequence of lattices sites $[m, X(m)],\ m = 0, 1, ..$ with $X(0) = 0$ and

$$\text{(3.7)}\qquad \begin{cases} X(m+1) & = \ X(m) + 1 \text{ with probability } p_u(X(m)) \ , \\ X(m+1) & = \ X(m) \text{ with probability } p_s(X(m)) \ , \\ X(m+1) & = \ X(m) - 1 \text{ with probability } p_d(X(m)) \ , \end{cases}$$

provided $|X(m)| < \min[m, J]$. To generate the random step in (3.7) by MC we use the generator "rand" in MATLAB, which generates a random number ξ with uniform distribution in the interval $0 < \xi < 1$. Then we set:

$$\text{(3.8)}\qquad \begin{cases} \text{if } \xi < p_u(X(m)) \text{ then} & X(m+1) \ = \ X(m) + 1 \ , \\ \text{else if } \xi < p_u(X(m)) + p_s(X(m)) \text{ then} & X(m+1) \ = \ X(m) \ , \\ \text{else} & X(m+1) \ = \ X(m) \ . \end{cases}$$

To calculate the value of the interest rate cap we generate a random path $[m, X(m)],\ 0 \le m \le M$, and calculate the value of the cap for this path just as in (3.2). Thus we set

$$\text{(3.9)}\qquad V = 0; \quad \text{for } m = M : -1 : 1$$

$$V \;=[\{r^{m-1}_{X(m-1)}-K\}^+\Delta t+V]\exp[-r^{m-1}_{X(m-1)}\Delta t]\ ;$$

end.

If $V_n,\ n=1,..,N$, are the values of V obtained in (3.9) from N independent MC simulations then

$$\text{MC value of interest rate cap} \ \simeq\ [V_1+\cdots+V_N]/N\ . \tag{3.10}$$

4. Black-Derman-Toys (BDT) Model

The BDT model is another short rate model which has some advantages and some disadvantages compared to the HW model. In HW we began with the SDE (2.1) and then derived the discrete model on a lattice. Here we shall go in the opposite direction, first defining the discrete model and then obtaining the corresponding SDE by taking the limit of the discrete model as $\Delta t\to 0$.

The lattice sites are points (m,j) in the integer lattice with $m=0,1,2,..$, and $0\le j\le m$. We associate with each lattice site (m,j) a time $t=m\Delta t$ and a short term rate r^m_j given by

(4.1)

$$r^m_j \ =\ r^m_0\exp\left[2j\beta^m\sqrt{\Delta t}\ \right]\ , \quad \text{where } r^m_0,\beta^m \text{ are parameters depending only on } m.$$

The parameters (r^m_0,β^m) are both required to be positive so $r^m_0>0$ is the lowest allowable short term rate at time $t=m\Delta t$. Transition probabilities in this model are simply given by the formulae

$$\begin{cases}(m,j)\to(m+1,j+1) & \text{with probability } \ 1/2\ ,\\ (m,j)\to(m+1,j) & \text{with probability } \ 1/2\ .\end{cases} \tag{4.2}$$

It is easy to see that the $\beta^m,\ m=1,2,..$, are volatilities. To see this we consider the increment $\log r(t+\Delta t)-\log r(t)$ in the logarithm of the short term rate during one time step. Thus from (4.1), (4.2)

$$\log r(m\Delta t)-\log r((m-1)\Delta t)) \ =\ \begin{cases}k \text{ with probability } 1/2,\\ k+2\beta^m\sqrt{\Delta t} \text{ with probability } 1/2,\end{cases} \tag{4.3}$$

where k is given by

(4.4)

$$k \ =\ \log r^m_j-\log r^{m-1}_j \ =\ \log r^m_0-\left(\frac{\beta^m}{\beta^{m-1}}\right)\log r^{m-1}_0+\left(\frac{\beta^m-\beta^{m-1}}{\beta^{m-1}}\right)\log r((m-1)\Delta t)\ .$$

Hence the random variable $\log r(t+\Delta t)-\log r(t)$ conditioned on $r(t)$ is Bernoulli with variance given by the formula

$$\mathrm{Var}\left[\log r(m\Delta t)-\log r((m-1)\Delta t))\ \middle|\ r((m-1)\Delta t)\ \right] \ =\ (\beta^m)^2\Delta t\ . \tag{4.5}$$

The formula (4.5) shows that β^m is the volatility of the increment in the logarithm of the short rate at time $t=(m-1)\Delta t$.

To calibrate the model to market data we proceed similarly to the HW model. Thus we define AD securities $Q^m_j,\ 0\le j\le m$, as in (2.39) by

$$Q^m_j \ =\ E\left[\exp\left\{-\int_0^{m\Delta t} r(t)\ dt\right\}\ ;\ r(m\Delta t)=r^m_j\ \right]\ . \tag{4.6}$$

The bond pricing equation analogous to (2.40) is now

$$P^{m+1} \ = \ \sum_{j=0}^{m} Q_j^m \exp[-r_j^m \Delta t] \ . \tag{4.7}$$

The recurrence relation for the AD securities analogous to (2.42) is given by

$$Q_j^{m+1} \ = \ \frac{1}{2} Q_j^m \exp[-r_j^m \Delta t] + \frac{1}{2} Q_{j-1}^m \exp[-r_{j-1}^m \Delta t] \quad \text{if } 0 < j < m+1. \tag{4.8}$$

At the boundaries $j = 0, m+1$ we have that

$$Q_0^{m+1} \ = \ \frac{1}{2} Q_0^m \exp[-r_0^m \Delta t] \ , \quad Q_{m+1}^{m+1} \ = \ \frac{1}{2} Q_m^m \exp[-r_m^m \Delta t] \ . \tag{4.9}$$

Evidently the Q_j^m, $0 \le j \le m$, are computable in terms of the parameter values $(r_0^{m'}, \beta^{m'})$, $0 \le m' < m$. Then (4.7) gives us one equation to determine the two parameter values (r_0^m, β^m). We clearly need a second equation, which should be obtained from market volatility information.

We can get an *approximate* second equation by making a comparison to Black's model for caplets. To see this we observe that

$$\log r(m\Delta t) \ = \ \log r_0^m + 2\beta^m \sqrt{\Delta t} \ X^m \ , \tag{4.10}$$

where X^m is the binomial variable for the sum of independent tosses of a fair coin. Thus

(4.11)

$$X^m \ = \ \sum_{k=1}^{m} Y_k \quad \text{where the } Y_k \text{ are independent and } Y_k = 0 \text{ or } 1 \text{ with probability } 1/2 \, .$$

From the central limit theorem the variable X^m is for large m approximately Gaussian with mean $m/2$ and variance $m/4$. We conclude that if $t = m\Delta t$ then

(4.12)

$$r(t) \ \simeq \ \text{exponential of a Gaussian variable with variance } (2\beta^m \sqrt{\Delta t})^2 m/4 \ = \ (\beta^m)^2 t \, .$$

Now we see from (1.7) that in Black's model the value of a caplet for the period $t \to t + \Delta t$ is given by

(4.13)

$$\text{value of caplet for the period } t \to t+\Delta t \ = \ P(0, t+\Delta t)\Delta t E\left[\max\{L(t) - K, 0\}\ \right] \ ,$$

where $L(t)$ is the exponential of a Gaussian variable satisfying

$$E[\ L(t)\ ] = F(0,t), \quad \text{variance of } \log L(t) \ = \ \sigma^2(t) t \ . \tag{4.14}$$

In (4.14) the quantity $F(0,t)$ is the instantaneous forward rate at time t defined in terms of the forward rates (1.3) by

$$F(0,t) \ = \ \lim_{\Delta t \to 0} F(0, t, t+\Delta t) \ . \tag{4.15}$$

The quantity $\sigma(t)$ is the instantaneous Black caplet volatility at time t, which is the volatility for the caplet over an infinitesimal time interval $(t, t+\Delta t)$.

Now from (3.2) the value of the caplet in (4.13) is given by

(4.16)

$$\text{value of caplet for the period } t \to t+\Delta t \ = \ E\left[\ \exp\left\{ -\int_0^t r(s)\ ds \right\} \max\{r(t) - K, 0\}\Delta t \ \ \right] \ .$$

If we make the approximation as in (3.6) that

$$(4.17)\quad E\left[\ \exp\left\{-\int_0^t r(s)\ ds\right\}\max\{r(t)-K,0\}\ \right]$$

$$\simeq\ E\left[\ \exp\left\{-\int_0^t r(s)\ ds\right\}\ \right]\ E\left[\ \max\{r(t)-K,0\}\ \right]\ =\ P(0,t)\,E\left[\ \max\{r(t)-K,0\}\ \right]\ ,$$

then (4.16) yields

(4.18)

$$\text{value of caplet for the period } t\to t+\Delta t\ =\ P(0,t)\Delta t E\left[\ \max\{r(t)-K,0\}\ \right]\ .$$

We can directly compare (4.13) and (4.18) since both $L(t)$ and $r(t)$ are exponentials of Gaussian variables. We can also write

$$(4.19)\qquad E[\ r(t)\ ]\ \simeq\ F(0,t)\quad\text{if yield curve volatility is low,}$$

since there is equality in (4.19) if yield curve volatility is zero. Now the LHS of (4.13), (4.18) are the same. Hence it follows from (4.12), (4.14) and (4.19) that setting $\beta^m=\sigma(m\Delta t),\ m=1,2,..$, is a reasonable assumption provided yield curve volatility is low. We can test whether our assumption (4.19) is *self-consistent* by evaluating $E[\ r(t)\ ]$ after calibration based on setting the β^m to equal Black implied volatilities for caplets, and then comparing with the RHS of (4.19). From (4.10), (4.11) we have that

$$(4.20)\quad E[\ r(t)\ ]\ =\ r_0^mE\left[\ \exp\{2\beta^m\sqrt{\Delta t}\ X^m\}\ \right]\ =\ r_0^mE\left[\ \exp\{2\beta^m\sqrt{\Delta t}\ Y_1\}\ \right]^m$$

$$=\ r_0^m\left[\ \frac{1}{2}+\frac{1}{2}\exp\{2\beta^m\sqrt{\Delta t}\}\ \right]^m\qquad\text{since the variable } Y_1 \text{ is Bernoulli.}$$

Having set the β^m to be Black implied volatilities for caplets, we can complete the calibration of the BDT model by solving (4.7) for r_0^m. Note that (4.7) is an *implicit* equation in x of the form

$$(4.21)\quad P^{m+1}\ =\ \sum_{j=0}^m a_j\exp[-b_jx]\ ,\quad\text{where the } a_j,b_j,\ 0\le j\le m,\ \text{are known.}$$

It can be solved efficiently by an iteration method such as Newton's method. In MATLAB one can use the equation solver "fsolve". Once the model has been calibrated the values of interest rate derivatives can be computed by the methods given in §3.

Finally we wish to find the continuum limit $\Delta t\to 0$ of the BDT model. From (4.3), (4.4) we have on setting $\beta^{m'}=\sigma(m'\Delta t),\ m'=1,2,..$, and $t=(m-1)\Delta t$ that

(4.22)

$$\log r(t+\Delta t)-\log r(t)\ \simeq\ \left[\frac{\sigma(t+\Delta t)-\sigma(t)}{\sigma(t)\Delta t}\log r(t)+\theta(t)\ \right]\Delta t+\sigma(t)[B(t+\Delta t)-B(t)]\ ,$$

where $B(\cdot)$ is Brownian motion. Note from (4.4), (4.5) that the mean and variance of the LHS and RHS of (4.22) agree to leading order in Δt. The function $\theta(t)$ in (4.22) refers to the terms in (4.4) which are deterministic and depend only on m. If we let $\Delta t\to 0$ in (4.22) we obtain the SDE

$$(4.23)\qquad d[\ \log r(t)\ ]\ =\ \left[\frac{\sigma'(t)}{\sigma(t)}\log r(t)+\theta(t)\ \right]dt+\sigma(t)\ dB(t)\ .$$

Thus in the BDT model it is the logarithm of the short rate which satisfies a linear SDE, whereas in the HW model the short rate itself satisfies a linear SDE.

We can easily explicitly solve (4.23). Multiplying (4.23) by $\sigma(t)^{-1}$ and setting $X(t) = \log r(t)$ we have from (4.23) that

$$\frac{dX(t)}{\sigma(t)} - \frac{\sigma'(t)}{\sigma(t)^2} X(t)\ dt \ = \ \frac{\theta(t)}{\sigma(t)}\ dt + dB(t)\ . \tag{4.24}$$

We can rewrite (4.24) as

$$d\left[\ \frac{X(t)}{\sigma(t)}\ \right] \ = \ \frac{\theta(t)}{\sigma(t)}\ dt + dB(t)\ , \tag{4.25}$$

and integrate. Thus we have that

$$\frac{X(t)}{\sigma(t)} \ = \ \frac{X(0)}{\sigma(0)} + \int_0^t \frac{\theta(s)}{\sigma(s)}\ ds + B(t)\ . \tag{4.26}$$

Hence $\log r(t)$ is a Gaussian variable with

$$\text{mean} \ = \ \frac{\sigma(t) \log r(0)}{\sigma(0)} + \sigma(t) \int_0^t \frac{\theta(s)}{\sigma(s)}\ ds\ , \quad \text{variance} \ = \ \sigma^2(t)t. \tag{4.27}$$

We have already mentioned that an advantage of the BDT model over HW is that the short rate is always positive. There are also disadvantages however. One rather subtle disadvantage is that the model cannot be mean reverting. To see why this is the case observe from (4.23) that for the model to be mean reverting the function $\sigma(\cdot)$ must have the property that there exists $a > 0$ and

$$\sigma'(t)/\sigma(t) \ \le -a \quad \text{for } t \ge 0. \tag{4.28}$$

We can rewrite (4.28) as

$$\frac{d}{dt}[\log \sigma(t) + at] \ \le \ 0 \quad \text{implies } \sigma(t) \ \le \ \sigma(0)e^{-at} \text{ for } t \ge 0. \tag{4.29}$$

Hence $\sigma(t)$ converges exponentially to 0 and since the variance of $\log r(t)$ is $\sigma^2(t)t$ this means the variance of $\log r(t)$ converges to zero. We conclude that the range of values taken by $r(t)$ in the BDT model as t becomes large will be wider than in the mean reverting HW model.

Finally we summarize some advantages and some disadvantages of the BDT model.

Advantage: Short term rate is always positive.

Advantage: Inputting market volatility data straightforward by using Black implied volatilities.

Disadvantage: Calibration to zero yield curve not explicit (need for equation solvers).

Disadvantage: No mean reversion of the short rate.

5. Forward Rate Models

University of Michigan, Department of Mathematics, Ann Arbor, MI 48109-1109
E-mail address: conlon@umich.edu