The General Hull-White Model and Super Calibration

John Hull and Alan White

Joseph L. Rotman School of Management
University of Toronto
105 St George Street
Toronto, Ontario M5S 3E6
Canada

Tel: (416) 978 8615 (Hull)
Tel: (416) 978 3689 (White)

August, 2000

The General Hull-White Model and Super Calibration

There are two major approaches to modeling the term structure of interest rates. One approach is to model the evolution of either forward rates or discount bond prices. This approach was first developed by Heath, Jarrow and Morton (HJM, 1992). In this paper they model the behavior of instantaneous forward rates. The method is both powerful (it contains many other term structure models as special cases) and easy to understand. It exactly fits the initial term structure of interest rates, it permits as complex a volatility structure as desired, and it can readily be extended to as many sources of risk as desired.

More recently the HJM model has been modified by Brace, Gatarek and Musiella (1997), Jamshidian (1997), and Miltersen, Sandmann, and Sondermann (1997) to apply to non-instantaneous forward rates. This modification has come to be known as the Libor Market Model (LMM). In one version, 3-month forward rates are modeled. This allows the model to exactly replicate observed cap prices that depend on 3-month forward rates. In another version forward swap rates are modeled. This allows the model to exactly replicate observed European swap option prices. The main difficulty with the HJM – LMM models is that they are difficult to implement by any means other then Monte Carlo simulation. As a result they are computationally slow and difficult to use for American or Bermudan style options.

The other major approach to modeling the term structure is to describe the evolution of the instantaneous rate of interest, the rate that applies over the next short interval of time. Short rate models are often more difficult to understand than models of the forward rate. However, they are implemented in the form of a recombining tree similar to the stock price tree first developed by Cox, Ross, and Rubinstein (1979). This makes them computationally fast and useful for valuing all types of interest-rate derivatives.

The Generalized Model

The generalized Hull-White model is a model in which some function of the short-rate obeys a Gaussian diffusion process of the following form

$$df(r) = \left[q(t) - a(t)\, f(r)\right] dt + s(t)\, dz \tag{1}$$

The function $q(t)$ is selected so that the model fits the initial term structure. The functions $a(t)$ and $s(t)$ are volatility parameters that are chosen to fit the market prices of a set of actively traded interest-rate options.

The generalized Hull-White model contains many popular term structure models as special cases. When $f(r) = r$, $a(t) = 0$ and s is constant it is the Ho-Lee (1986) model. When $f(r) = r$ and $a(t)$ is not zero it is the original Hull-White (1990) model. In both these models future interest rates of all maturities are normally distributed and there are many analytic solutions for the prices of bonds and options on bonds. When $f(r) = \sqrt{r}$ it is a model developed by Pelsser (1996) and when $f(r) = \ln r$ it is the Black-Karasinski (1991) model which is perhaps the most popular version currently in use. In this model the future short-rate is log-normally distributed and rates of all other maturities are approximately log normally distributed.

In the next section of the paper we will describe how this class of models is implemented using a recombining trinomial tree. In the section on calibration we will discuss how the model parameters are chosen and finally, in the section on super-calibration we will show how the functional form $f(r)$ can be selected.

Implementation

In this section we will describe how the generalized model is implemented in a recombining trinomial tree. Initially we will assume that the volatility parameters, $a(t)$ and $s(t)$, and the functional form, $f(r)$, have been selected. Later we describe how these are chosen.

First we set the current time to 0 and define a deterministic function g, which satisfies

$$dg = \left[q(t) - a(t)\,g(t)\right]dt$$

We then define a new variable, x, that is

$$x(r,t) = f(r) - g(t)$$

The new variable obeys a much simpler diffusion process

$$dx = -a(t)\,x\,dt + s(t)\,dz$$

The initial value of g is chosen so that the initial value of x is 0.[1] This process is mean reverting to 0 so that if x starts at 0 the unconditional expected value of x at all future times is 0.

---

[1] When the reversion rate is constant the form of g is $g(t) = g(0)e^{-at} + \int_0^t q(s)\,e^{-a(t-s)}ds$. While this looks ominous we do not actually ever have to determine its exact form. The addition of this function to the process is just a device that makes the implementation simpler.

Building a tree for $f(r)$ involves 4 steps. The first step is to select the spacing of the tree nodes in the time dimension. The second step is to decide on the spacing of the nodes in the interest-rate dimension. The third step is to choose the branching process for $x(r, t)$ through the grid of nodes. Once this is complete, the fourth step involves shifting the tree by the value of g at each point in time. This then results in a tree for f.

1. Choosing the times at which nodes are placed

When a term structure model is implemented it is usually for some specific purpose such as pricing an option on a swap. As a result it is convenient to construct the tree in such a way that we have nodes on specified dates such as payment and exercise dates. Suppose we wish to build an n-step tree with nodes at times t_0, t_1, t_2, …, t_n where $t_0 = 0$, $t_i > t_{i-1}$ and $t_n = T$, the longest date to be considered. Since the values of all bonds, swaps and other instruments are computed by discounting their payoffs back through the tree, T must be chosen so that no payments occur after T. We should also ensure that we have chosen our node times, t_i, so that there is a set of nodes on every payment date. Other node times can be selected to increase the resolution of the tree.

2. Choosing the values of x where nodes are to be placed

Once the times at which nodes are to be placed have been chosen, at each time step we must choose the values of x where nodes are to be placed. First we place a node at $x = 0$ at each time step. Then at each time step t_i ($i = 1, \ldots, n$) we place nodes at $\pm\Delta x_i$, $\pm 2\Delta x_i$, …, $\pm m_i \Delta x_i$. The determination of the value of m_i will be explained in the following section. In choosing the Δx_i, the only constraint we face is that the spacing of the nodes must be wide enough to represent the volatility of x at that time. This is achieved by setting the x-spacing at time t_i to[2]

$$\Delta x_i = s\,(t_{i-1})\sqrt{3(t_i - t_{i-1})} \tag{2}$$

The next stage of the implementation is to determine how the nodes in (x, t) space will be connected together. This will also determine the m_i's, the indices of the highest and lowest nodes that are attainable at each time step.

3. Choosing the branching process

We choose the branching through the tree so that at every point in the tree we are mimicking the diffusion process as closely as possible. This is done by ensuring that the expected change and the variance of the change in x seen on the tree are the same as predicted by the diffusion process for x. At each node in the tree we select the branching process and the branching probabilities accordingly.

Suppose that we are at some node $j\Delta x_i$ at step i and propose to branch to nodes $(k-1)\Delta x_{i+1}$, $k\Delta x_{i+1}$, and $(k+1)\Delta x_{i+1}$ at step i+1. From the diffusion process for x we calculate the expected mean change in x over the next time interval, $E(dx) = M$, and the second moment of x, $E(dx^2) = V + M^2$.[3] Let the probability of branching to $(k-1)\Delta x_{i+1}$,

[2] The node spacing can be set to $\Delta x_i = s\,(t_{i-1})\sqrt{n(t_i - t_{i-1})}$ for a range of values of n without impairing the numerical procedure. The choice $n = 3$ is made because this allows the numerical procedure to exactly replicate the first 5 moments of the distribution of $x(t_i)|x(t_{i-1})$ when the reversion rate is zero. This produces a slightly more rapid convergence than do other values of n.

[3] A reasonable approximation is $M = -x\,a(t_i)(t_{i+1} - t_i) = -j\Delta x_i\, a(t_i)(t_{i+1} - t_i)$ and $V = s^2(t_i)(t_{i+1} - t_i)$. When a and s are constant more exact calculations are possible.

$k\Delta x_{i+1}$, and $(k+1)\Delta x_{i+1}$ be p_d, p_m, and p_u respectively. Matching the mean and variance gives

$$\begin{aligned} j\Delta x_i + M &= k\Delta x_{i+1} + (p_u - p_d)\Delta x_{i+1} \\ V + (j\Delta x_i + M)^2 &= k^2\Delta x_{i+1}^2 + 2k(p_u - p_d)\Delta x_{i+1}^2 + (p_u + p_d)\Delta x_{i+1}^2 \end{aligned} \tag{3}$$

Solving equation (3) we find that

$$\begin{aligned} p_u &= \frac{V}{2\Delta x_{i+1}^2} + \frac{a^2 + a}{2} \\ p_d &= \frac{V}{2\Delta x_{i+1}^2} + \frac{a^2 - a}{2} \\ p_m &= 1 - \frac{V}{\Delta x_{i+1}^2} - a^2 \end{aligned} \tag{4}$$

where

$$a = \frac{j\Delta x_i + M - k\Delta x_{i+1}}{\Delta x_{i+1}}$$

is the distance from the expected value of x to the central node to which we are branching. If $V = s^2(t_i)(t_{i+1} - t_i)$ and $\Delta x_{i+1} = s(t_i)\sqrt{3(t_{i+1} - t_i)}$ it can be shown that all the branching probabilities are positive if $-\sqrt{2/3} < a < \sqrt{2/3}$. That is, when branching from a point $j\Delta x_i$, we should choose as the central node of the 3 successor nodes a node within $\sqrt{2/3}\,\Delta x_{i+1}$ of the expected outcome. Usually we choose the node closest to the expected outcome by setting k to the value of $(j\Delta x_i + M)/\Delta x_{i+1}$ rounded to the nearest integer. This ensures we are within $\Delta x_{i+1}/2$ of the expected outcome and the condition for positive probabilities is satisfied.

The procedure we have just described determines the tree branches and branching probabilities. It also defines the highest and lowest possible node at each step. The highest node at step i+1, m_{i+1}, is determined by the branching from the highest node at step i, m_i. Similarly, the lowest node at step i+1, $-m_{i+1}$, is determined by the branching from the lowest node at step i, $-m_i$. Since at step 0 there is only one node $m_0 = 0$. From this the highest and lowest nodes at step 1 and all subsequent steps can be determined.

We illustrate the calculation with an extreme example. We suppose that t_0=0, t_1=1.5, t_2=1.6, and t_3=2.0 so that the time steps are of widely varying lengths. (In most applications they are much more equal than this.) We suppose that the volatility parameters are $a(t)$=1.0 and $\sigma(t)$=0.30 for all t. The node spacing at each step is determined using equation (2). This gives Dx_1=0.6364, Δx_2=0.1643, and Δx_3=0.3286. The grid of nodes on the tree is therefore as shown in Table 1.

The next step is to compute the branching process. Starting at the root node (t=0 and x=0), we compute $x+M=x-ax\times 1.5=0$ and $V = 0.30^2 \times 1.5 = 0.135$. The node closest to the expected outcome is the node k = 0 at t = 1.5. For this node a = 0 and using equation (4) the branching probabilities are p_d = 0.1667, p_m = 0.6667 and p_u = 0.1667. Similarly at the highest node at step 1 (t = 1.5 and x = 0.6364), $x + M = x - ax \times 0.1 = 0.5728$, $V = 0.30^2 \times 0.1 = 0.009$, $\left(x+M\right)/\Delta x_{i+1} = 3.486$ so k = 3, and $a = \left(0.5728 - 3\times 0.1643\right)/0.1643 = 0.4857$. The results for every node are in the Table 2 and the shape of the tree is shown in Figure 1.

4. Adjusting the tree

The final stage of the tree building process involves adding the function $g(t)$ to the value of x at each node. Since $g(t)$ is a function of $q(t)$ and the function $q(t)$ is selected so that the model fits the term structure, the de facto process is to adjust the nodes in the tree so that it correctly prices discount bonds of all maturities. This is done in a sequential process starting at the root node.

We denote node (i, j) as the node on the tree at time t_i for which $x=j\Delta x_i$ ($0 \le i \le n$; $-m_i \le j \le m_i$) and define

g_i: $g(t_i)$

x_{ij}: value of x at node (i, j)

f_{ij}: value of $f(r)$ at node (i, j). This is $x_{ij} + g_i$.

r_{ij}: interest rate at node (i, j) . This is $f^{-1}(x_{ij} + g_i)$

$Q(i, j|h, k)$: value at node (h, k) of a security that pays off \$1 at node (i, j) and nothing at any other node.[4]

$p(i, j|h, k)$: the probability of transiting from node (h, k) to node (i, j)

Q_{ij}: $Q(i, j|0, 0)$

The variable $Q(i, j|h, k)$ is known as an Arrow-Debreu (AD) price. We will refer to the Q_{ij} as the root AD price for node (i, j).

---

[4] The value of any security with deterministic payoffs can be easily computed using the AD prices. Letting C_{ij} be the payment received at the ij'th node the value of the security at the hk'th node is then

The root AD price for node (i, j) can be determined once the root AD prices for all nodes at time t_{i-1} have been determined. To see this we note that

$$Q(i, j \mid i-1, k) = p(i, j \mid i-1, k)\exp\left[-r_{i-1,k}\left(t_i - t_{i-1}\right)\right]$$

and

$$\begin{aligned} Q_{ij} &= \sum_k Q(i, j \mid i-1, k)\, Q_{i-1,k} \\ &= \sum_k p(i, j \mid i-1, k)\exp\left(-r_{i-1,k}\left(t_i - t_{i-1}\right)\right) Q_{i-1,k} \end{aligned} \qquad (5)$$

where the summation is over all nodes at time step t_{i-1}.

Now consider a discount bond that pays \$1 at every node at time step i+1. Let P_{i+1} be the price at node (0, 0) of this discount bond and let V_{ij} be the value of this bond at node (i, j). The process for determining the adjustment g_i at step i involves two stages. First we determine Q_{ij} for every node j at step i. Using these root AD prices we then compute the value of P_{i+1}. Since the discount bond pays \$1 at every node at t_{i+1} the value at the ij'th node is

$$\begin{aligned} V_{ij} &= \exp\left[-r_{i,j}\left(t_{i+1} - t_i\right)\right] \\ &= \exp\left[-f^1\left(x_{ij} + g_i\right)\left(t_{i+1} - t_i\right)\right] \end{aligned}$$

and the present value is

$\sum_{i>h}\sum_j Q(i, j \mid h, k)\, C_{ij}$ where the summation is taken over all time steps, i, later than h and all nodes at each time step, j.

$$
\begin{aligned}
P_{i+1} &= \sum_j Q_{ij} V_{ij} \\
&= \sum_j Q_{ij} \exp\left[-f^{-1}\left(x_{ij} + g_i\right)\left(t_{i+1} - t_i\right)\right]
\end{aligned} \tag{6}
$$

The value of g_i is adjusted until the value computed using equation (6) matches the price of the discount bond computed from the current term structure.

The implementation of this two-stage process proceeds in the following way. The value of a security that pays \$1 at the root node is \$1 so $Q_{00} = 1$. Based on the value of Q_{00} equation (6) is used to compute g_0 to match the price of a discount bond maturing at t_1. This allows us to use equation (5) to compute Q_{1j} for every node j, which then allows us to use equation (6) to compute g_1 and so on.

To complete the illustration of the tree-building process we will now fit our example tree to a term structure. Suppose that $x = f(r) = \ln r$ $\left(r = f^{-1}(x) = e^x\right)$ and that the term structure of continuously compounded discount bond yields is given in Table 3. The tree adjustment process is to first set $Q_{00} = 1$. Then solving equation (6) at the root node

$$
\begin{aligned}
P_1 &= Q_{00} \exp\left[-f^{-1}\left(x_{ij} + g_i\right)\left(t_{i+1} - t_i\right)\right] \\
0.9277 &= \exp\left[-\exp\left(0 + g_0\right)(1.5)\right]
\end{aligned}
$$

we find $g_0 = -2.9957$ and $r_{00} = f^{-1}\left(x_{00} + g_0\right) = \exp(-2.9957) = 0.05$. This rate is used to calculate

$$
\begin{aligned}
Q_{1,1} &= Q_{00} p_u \exp\left(-r_{00} \times 1.5\right) = 0.1546 \\
Q_{1,0} &= Q_{00} p_m \exp\left(-r_{00} \times 1.5\right) = 0.6185 \\
Q_{1,-1} &= Q_{00} p_d \exp\left(-r_{00} \times 1.5\right) = 0.1546
\end{aligned}
$$

where the probabilities, p_u, p_m and p_d, are the probabilities of transiting from the root node to the 3 nodes at time step 1. With this in hand equation (6) is used to find g_1 and so on. The results of the calculations are shown in the Table 4

This completes the construction of the tree for a log-normally distributed short-rate that exactly fits the term structure. It is worth noting at this point that the functional form, $f(r)$, only comes into play at the stage where the term structure is being fit (although as we will show, it does have an impact on the volatility parameters chosen). Prior to the term-structure fitting stage, the tree building process is completely generic. Also note that when the tree was being fit to the term-structure, in order to compute the interest rates at the fourth time step we had to specify a fifth time step at time 2.5 years. This was necessary to allow us to define the term of the rates that are being determined at the fourth step. In this case they are 0.5-year rates.

Calibration

Calibration is the process of determining the volatility parameters that are used in the term structure model. It is analogous to selecting the volatility that will be used when implementing the Black-Scholes model to price equity options. In the case of the generalized Hull-White model the volatility parameters that are to be chosen are the functions $a(t)$ and $s(t)$. The procedure is to choose the volatility parameters so that the tree implementation of the term structure model accurately replicates the market prices of actively traded options. Specifically we use a numerical procedure such as the Levenberg-Marquardt algorithm to find the set of volatility parameters that minimizes the sum of the squares of the differences between the model prices and market prices for these options.

Because the volatility parameters are functions it is necessary to parameterize them before starting the calibration process. Typically we approximate the volatility functions with piecewise linear functions. This corresponds to selecting a set of times T_0, T_1, T_2, …, T_m where $T_0 = 0$, $T_i > T_{i-1}$ and then defining the reversion rate function as

$$a(t) = a_i + b_i t \qquad T_i \le t < T_{i+1}$$

subject to

$$a_i + b_i T_{i+1} = a_{i+1} + b_{i+1} T_{i+1}, \quad b_0 = 0, \qquad b_m = 0$$

The first condition ensures that the function is continuous and the second and third ensure that it is constant in the first time interval and beyond the last specified date.[5] These constraints ensure that there are m degrees of freedom in the parameter set. The volatility function is defined in an analogous way as

$$s(t) = g_i + d_i t \qquad T_i \le t < T_{i+1}$$

subject to

$$g_i + d_i T_{i+1} = g_{i+1} + d_{i+1} T_{i+1}, \quad d_0 = 0, \qquad d_m = 0$$

The choice of the number of corner points in the volatility functions and at what times the corners should be placed is more of an art than a science. Using more corner points gives more degrees of freedom and permits a better fit to the observed market prices. Often the number and timing of the corner points are determined by the terms of the options that are used in the calibration. If we have m calibrating options with m distinct maturity dates

[5] Neither of these conditions is required. They are used only because of a belief that the volatility functions should be continuous and bounded. An alternative parameterization that seems to work well is a step function in which the parameters are piecewise constant. Note that the time divisions used for the two volatility functions do not need to be the same.

then holding one volatility function constant (usually the reversion rate) and choosing the corner points of the other to be the option maturity times ensures that we can fit the option prices exactly.

The most common source of option prices for calibration purposes are quotes that are available from brokers on European-style swap options and caps and floors. Table 5 shows a typical panel of USD swap option quotes for August 6, 1999. This table contains the volatilities for a range of at-the-money swap options. These are the volatilities that if used in the market standard Black's swap option-pricing model, result in the mid-market prices for the options. The market prices of the options range from $0.12 for the 30-day option on a $100 notional 1-year swap to $5.45 for the 5-year option on a 10-year swap.

The results of fitting both the normal and the lognormal versions of the model to this data using only a single reversion rate and a single volatility are shown in Table 6. This table shows the best-fit reversion rate, the best-fit volatility and the root mean square pricing error[6] (RMSE). The fit of the model to the option prices is moderately good for both versions of the model although the normal version fits somewhat better than the lognormal version. The mean absolute percentage pricing error (the average of the absolute price error divided by the market price) is about 2.5%. Those who are not familiar with the various forms of term structure models should also note that the magnitude of the volatility parameter is dependent on the functional form of the model. In the normal model the volatility parameter corresponds to the standard deviation of

---

[6] The root mean square error is defined as $\sqrt{\sum_{i=1}^{n}\left(P_{\text{model}} - P_{\text{market}}\right)^2 / n}$ where n is the number of option prices being fit.

annual changes in the short-term rate of interest while in the lognormal model it is the standard deviation of proportional changes in the rate. Thus, if interest rates are about 7% a 1.4% annual standard deviation roughly corresponds to an annual standard deviation of proportional changes of 20%.

To improve the fit we can use more volatility parameters. Table 7 shows the results of increasing the parameter set so that there is a corner in both the reversion rate and volatility functions at every option maturity date. Comparing Tables 6 and 7, we see that increasing the number of volatility parameter from 2 to 16 does improve the fit, but not dramatically so. The volatility parameter for the normal model is relatively constant and the reversion rate changes only five times suggesting that about the same fit could be achieved with far fewer parameters. In the lognormal model, by contrast, both $a(t)$ and $\sigma(t)$ are highly variable.

Some experimentation reveals that it is not possible to fit this full panel of option prices using our model or indeed any one-factor Markov model of the term structure. As a result when these types of models are used in practice they are calibrated in the same way that equity and F/X option pricing models are calibrated. A different volatility parameter set is used for every different option or for every different type of option. Usually the volatilities of the European options that are used to hedge the option in question will be used for calibration.

For example, a common use of these models is the pricing of Bermudan swap options. To calibrate our model to price Bermudan swap options we use a diagonal strip of volatilities from Table 5 for calibration. If we are interested in pricing a 5-year Bermudan swap option we note that if it is exercised at the 1-year point in its life it is similar to a 1-year

European option on a 4-year swap. Similarly exercise at the 2-year point is similar to a 2-year European option on a 3-year swap, and so on. As a result we use the 1x4, 2x3, 3x2 and 4x1 swap option volatilities to calibrate the model and we will likely use these options to hedge the Bermudan option. By using 4 volatility parameters we can exactly fit the calibrating option prices with our model and achieve a good hedge – or at least a good hedge for the prices calculated by the model.

Super Calibration

In the previous section we discussed how the volatility parameters for a particular form of the model could be determined from market prices of options. In this section we describe how the functional form of the model can also be determined from the market prices of options.

Black's model, the market standard for caps and European swap options, assumes that interest rates are lognormally distributed. If rates really were lognormally distributed, the volatility used to price a cap or a swap option would be independent of the option strike rate. In the last year the USD cap market has developed to the point that brokers are now able to provide volatility quotes for in- and out-of-the-money caps and floors. The usual practice is to provide at-the-money volatility quotes for the standard set of caps and to provide a table of spreads to be added to the volatilities of in- and out-of-the-money caps. A typical set of broker quotes for July 27, 1999 is shown in Table 8.

Since the market volatilities for caps and floors are not independent of their strike rates we can conclude that the lognormal assumption does not reflect the market perception of the distribution of rates. Table 8 shows that volatilities for in-the-money caps are significantly higher than those for at-the-money caps. Except for very long maturities, out-of-the-money caps also have somewhat higher volatilities that at-the-money caps. The market's perception is therefore that very low rates and (to a lesser extent) very high rates are more likely than the lognormal distribution would suggest.

The term structure models implied by equation (1) assumes that some function of the short rate, $x = f(r)$, follows a normal mean-reverting process. To understand the role that the functional form, *f*(*r*), plays note that the process that the short rate, *r*, obeys is

$$dr = \cdots dt + \frac{\partial h(x)}{\partial x}\sigma(t)dz \tag{7}$$

where *h* is the inverse of the function *f*, that is, $r = h(x)$ The primary effect of the choice of the functional form is in its impact on the volatility component of this process, $\sigma(t)\partial h(x)/\partial x$. This determines the relation between the level of rates and the variability of rates. We now propose a more general model in which $\sigma(t)\partial h(x)/\partial x = \sigma(t)s(r)$ for some function of the level of rates, $s(r)$. The function $\sigma(t)s(r)$ is known as the local standard deviation of the rate and $\sigma(t)s(r)/r$ is the local volatility. In this paper we have so far considered two cases:

- $x = f(r) = r$ or $r = h(x) = x$ for which $\sigma(t)s(r) = \sigma(t)$, rates always have the same level of variability and future rates are normally distributed. This is the original Hull-White model.

- $x = \ln(r)$ or $r = \exp(x)$ for which $s(t)s(r) = s(t)r$, the variability of rates is proportional to the level of rates and rates are log normally distributed. This is the Black-Karasinski model.

These two models have $s(r) = 1$ and $s(r) = r$. Just as the volatility functions, $a(t)$ and $s(t)$, are constructed as piecewise linear functions, $s(r)$ can also be constructed as a piecewise linear function. This is done by selecting a number of different rates, $r_i > 0$ for $i = 1, 2, \ldots, n$, and the corresponding values of $s(r) > 0$, s_i for $i = 1, 2, \ldots, n$. We usually force $s(r)$ to pass through the origin. This ensures that as r becomes small the variability of rates vanishes and negative rates do not occur. The form of $s(r)/r$ for the three models is shown in Figure 2.

The selection of the values of s_i for $i = 1, 2, \ldots, n$ now becomes part of the calibration exercise. We choose the values that result in a term-structure model implementation that most closely replicates the market prices of the options. Our least squares best fit criterion is the same as before. Since the variability of the short rate in equation (7) is $s(t)s(r)$ it is not possible to determine the forms of both $s(t)$ and $s(r)$ simultaneously. As a result, we first find the $s(t)$ that best fits the at-the-money options and then, holding that fixed, find the $s(r)$ that best fits the prices of the in- and out-of-the-money options.

To illustrate the effect of calibrating the functional form to the volatility of in- and out-of-the-money options we set $s(t) = 1$ and find the best $s(r)$ to fit the prices of 3-year caps and floors. The corner points of $s(r)$ are set at the at-the-money rate ±0.5%, ±1%, and

±2%. This process is then repeated for the 7-year and 10-year caps and floors. The best-fit functional form of the local volatility for each of the 3 maturities is shown in Figure 3. The overall result shown in Figure 3 is not surprising. In order to raise the price (and implied volatility) of in- and out-of-the-money caps and floors we have to increase the local volatility as we move away from the money. The shorter the life of the option, the more extreme the adjustment becomes.

Conclusion

In this paper we have explained how a general model of the short-rate can be implemented and calibrated to market data. The calibration process includes the selection of the functional form of the term structure model that best fits the prices of in- and out-of-the-money options. Although not discussed in this paper, the super calibration process is also useful in economies like Japan's where interest rates are very low. In this situation if a normal model is used the probability of rates becoming negative is very large, while if a lognormal model is used the volatilities must be in excess of 100% to capture the observed variability of rates. A lognormal model with these large volatilities implies that rates will become extremely variable when they rise above 1%. This issue is discussed in more detail in Hull and White (1997).

The super calibration procedure described in this paper is in the same spirit as the implied tree methodology for equity options developed by Derman, Kani and Chriss (1996), and Rubinstein (1994). These authors made the local volatility of the stock price a function of time and the stock price and developed procedures to infer the local volatility from option prices. The super calibration procedure also suffers from the same weakness as the implied tree methodology, which is that we are adding many free parameters to our model in an attempt to force it to fit a complex data set. This does not result in a model that more accurately reflects the way the term structure actually evolves. It is a model that better reproduces observed market prices.

There is a range of views on what is best in fitting a model to data. At one extreme is what we might call the academic's view that simple, stationary models are best. This means that the volatility parameters should not be functions of time and that the functional form of the model should not change over time. The behavior of models with these properties will be the same in the future as it is now. However, if we restrict ourselves to stationary models we can only approximately fit observed market prices. At the other extreme is what we might call the trader's view that the model should exactly fit all observed option prices. If this is done many free volatility parameters must be estimated and the model becomes highly non-stationary. The future behavior of the model may be very different from its current characteristics. In particular the future option volatilities implied by the model may be very different from the volatilities seen today. Our view is that a moderate approach should be taken in fitting a model to observed option prices. Modest non-stationarity does not seriously affect the future behavior of the model and allows a good fit to today's prices.

References

Black, F. and P. Karasinski 1991. "Bond and Option Pricing When Short Rates Are Lognormal," *Financial Analysts Journal*, vol.47, no. 4, (July/August 1991):52–59.

Brace, A., D. Gatarek, and Musiela 1997. "The Market Model of Interest Rate Dynamics," *Mathematical Finance*, vol.7, no. 2, (April):127–55.

Cox, J., S. Ross, and M. Rubinstein 1979. "Option Pricing: A Simplified Approach," *Journal of Financial Economics*, vol. 7, no. 3, (September):229–64.

Derman, E., I. Kani, and N. Chriss 1996. "Implied Trinomial Trees of the Volatility Smile," *Journal of Derivatives*, vol. 3, no. 4, (Summer):7–22.

Heath, D., R. Jarrow, and A. Morton 1992. "Bond Pricing and the Term Structure of Interest-Rates: A New Methodology," *Econometrica*, vol. 60, no. 1, (January):77–105.

Ho, T. S. Y., and S.-B. Lee 1986. "Term Structure Movements and Pricing Interest Rate Contingent Claims," *Journal of Finance*, vol. 41, no. 5, (December):1011–29.

Hull, J., and A. White 1990. "Pricing Interest Rate Derivatives Securities," *The Review of Financial Studies*, vol. 3, no. 4, 573–92.

Hull, J., and A. White 1997. "Taking Rates to the Limit," *RISK*, (December):168–9.

Jamshidian, F. 1997. "LIBOR and Swap Market Models and Measures," Finance and Stochastics, vol.1, no. 4, 293–330.

Miltersen, K., K. Sandmann, and D. Sondermann 1997. " Closed Form Solutions for Term Structure Derivatives with Log Normal Interest Rates," *Journal of Finance*, vol. 52, no.1 (March):409–30.

Pelsser, A. A. J. 1996. "Efficient Methods for Valuing and Managing Interest Rate and Other Derivative Securities," Ph.D. dissertation, Erasmus University, Rotterdam.

Rubinstein, M. 1994. "Implied Binomial Trees," *Journal of Finance*, vol 49, no.3, (July):771–818.

Table 1 Grid of Tree Nodes			
t = 0	t = 1.5	t = 1.6	t = 2.0
	…	…	…
	…	0.3286	0.6573
	0.6364	0.1643	0.3286
0.0000	0.0000	0.0000	0.0000
	–0.6364	–0.1643	–0.3286
	…	–0.3286	–0.6573
	…	…	…

Table 2
Tree branching calculations

t	x	M	V	k	a	p_u	p_m	p_d
0	0	0	0.135	0	0.0000	0.1667	0.6667	0.1667
1.5	0.6364	-0.0636	0.009	3	0.4857	0.5275	0.4308	0.0418
1.5	0.0000	0.0000	0.009	0	0.0000	0.1667	0.6667	0.1667
1.5	-0.6364	0.0636	0.009	-3	-0.4857	0.0418	0.4308	0.5275
1.6	0.6573	-0.2629	0.036	1	0.2000	0.2867	0.6267	0.0867
1.6	0.4930	-0.1972	0.036	1	-0.1000	0.1217	0.6567	0.2217
1.6	0.3286	-0.1315	0.036	1	-0.4000	0.0467	0.5067	0.4467
1.6	0.1643	-0.0657	0.036	0	0.3000	0.3617	0.5767	0.0617
1.6	0.0000	0.0000	0.036	0	0.0000	0.1667	0.6667	0.1667
1.6	-0.1643	0.0657	0.036	0	-0.3000	0.0617	0.5767	0.3617
1.6	-0.3286	0.1315	0.036	-1	0.4000	0.4467	0.5067	0.0467
1.6	-0.4930	0.1972	0.036	-1	0.1000	0.2217	0.6567	0.1217
1.6	-0.6573	0.2629	0.036	-1	-0.2000	0.0867	0.6267	0.2867

Table 3		
The term structure of continuously compounded discount bond yields		
Time to Maturity	Yield	Bond Price
1.5	5.00%	0.9277
1.6	5.10%	0.9216
2.0	5.25%	0.9003
2.5	5.30%	0.8759

Table 4
Fitting the tree to the term structure

r_{ij} (%)				Q_{ij}				V_{ij}			
		10.664				0.0806				0.9582	
		9.048				0.0658				0.9645	
		7.677	10.238			0.0064	0.0302			0.9698	0.9501
	11.663	6.514	7.370		0.1546	0.1024	0.2023		0.9884	0.9743	0.9638
5.000	6.172	5.527	5.306	1.0000	0.6185	0.4098	0.4306	0.9277	0.9938	0.9781	0.9738
	3.266	4.689	3.820		0.1546	0.1024	0.2059		0.9967	0.9814	0.9811
		3.979	2.750			0.0064	0.0313			0.9842	0.9863
		3.376				0.0664				0.9866	
		2.864				0.0813				0.9886	

g0	g1	g2	g3
-2.9957	-2.7851	-2.8956	-2.9364

Table 5

Mid-market volatilities for at the money swap options. The swap is assumed to start at the expiry of the option so the total life of the transaction is the sum of the option life and the swap life.

	Swap Life (Years)						
Option Life	1	2	3	4	5	7	10
30-day	19.00	19.50	19.50	19.50	19.50	19.50	19.50
3-month	19.50	20.13	20.13	20.13	19.98	19.98	19.98
6-month	19.90	19.75	19.75	19.70	19.60	19.50	19.50
1-year	21.55	20.80	20.20	19.90	19.60	19.20	18.78
2-year	21.30	20.40	19.85	19.30	19.00	18.70	18.20
3-year	20.80	19.75	19.20	18.85	18.60	18.20	17.63
4-year	20.43	19.20	18.80	18.40	18.10	17.60	17.03
5-year	19.85	18.73	18.28	17.93	17.58	16.98	16.43

Table 6

Best fit volatility parameters for the normal and log-normal version of the model.

Model	Reversion rate, a	Volatility, s	RMSE
Normal	0.0267	0.0146	0.0564
Lognormal	0.0243	0.2093	0.0745

Table 7

Best fit volatility parameters for the normal and lognormal versions of the model

	Normal		Lognormal	
	$a(t)$	$s(t)$	$a(t)$	$s(t)$
05-Sep-99	0.1878	0.0147	0.0487	0.2144
05-Nov-99	0.0205	0.0135	0.0596	0.2137
04-Feb-00	0.0010	0.0135	0.0007	0.1669
05-Aug-00	0.0010	0.0136	0.0002	0.2261
05-Aug-01	0.0003	0.0133	0.0005	0.1513
05-Aug-02	0.0003	0.0132	0.0002	0.2199
05-Aug-03	0.0010	0.0130	0.0006	0.1436
04-Aug-04	0.0212	0.0130	0.0140	0.2071
RMSE	0.0310		0.0292	

Table 8

Volatility adjustments for in- and our-of-the-money caps and floors for July 27, 1999.

		Cap Strike – At-The-Money Strike (%)							
Cap Life	ATM Vols	-3	-2	-1	-0.5	0.5	1	2	3
1-year	14.88	-	-	1.00	0.50	0.00	1.00	-	-
2-year	18.38	3.00	2.00	1.00	0.50	0.50	1.00	1.25	1.50
3-year	19.19	3.15	2.15	1.15	0.75	0.70	0.75	1.10	1.10
4-year	19.50	3.50	2.50	1.50	0.75	0.50	0.50	1.00	1.00
5-year	19.50	3.00	2.00	1.20	0.80	0.00	0.50	1.00	1.00
7-year	18.88	3.00	2.00	1.00	0.50	0.00	0.00	0.00	0.00
10-year	18.19	3.00	2.00	1.00	0.50	0.00	-0.25	-0.50	-0.50

Figure 1
Tree Branching Structure

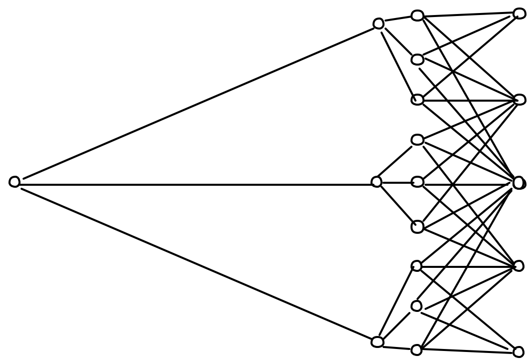

Figure 2
The relation between the level of rates and local volatility

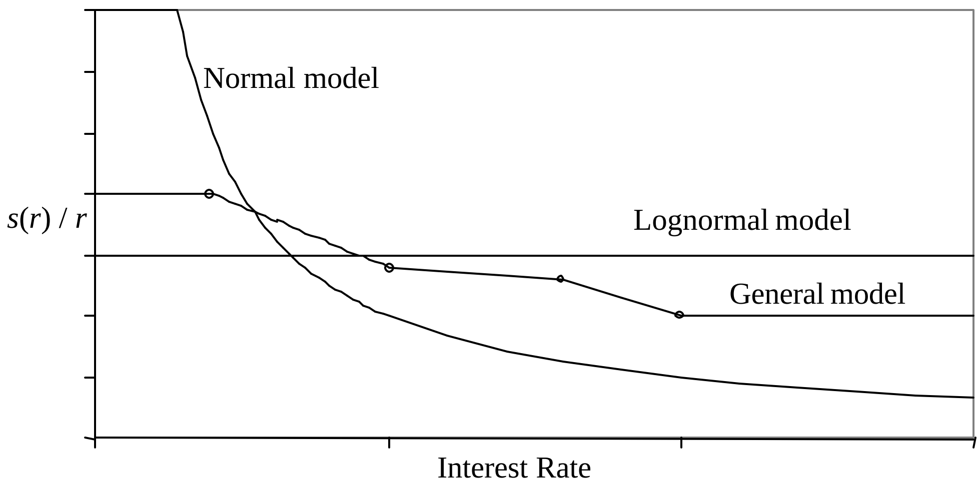

Figure 3
Best-fit Local Volatility