

TEACHING NOTE 02-01: THE HEATH-JARROW-MORTON TERM STRUCTURE MODEL

Version date: August 19, 2008

c:\Class Material\Teaching Notes\TN02-01.doc

One of the most significant discoveries in modern financial research is the Heath-Jarrow-Morton term structure model. It has revolutionized the theory and practice related to interest rate derivatives as well as interest rate products in general. The Heath-Jarrow-Morton model, or HJM as it is commonly known, is at the head of a family tree of models that take the initial term structure as given, affix the conditions required to preclude arbitrage, and produce trading strategies that lead to the valuation of all interest-dependent claims. Like its predecessor Ho-Lee, HJM is considered an arbitrage-free model but not a general equilibrium model in the genre of Vasicek or Cox-Ingersoll-Ross. Those models derive market equilibrium conditions leading to the prices of securities, but they do not fit the current term structure. Hence, they can admit arbitrage if used in current markets. By accommodating the existing term structure, HJM, as well as Ho-Lee, can be used in current markets.

But what sets HJM apart from Ho-Lee is the fact that it admits a wide range of structures for the sources driving interest rates. We should think of an interest rate as being driven by one or more “factors”, which are sources of noise or uncertainty. While Ho-Lee is a one-factor model, HJM can accommodate any finite number of factors, though with each factor there is a considerable increase in complexity and practicality. In addition HJM admits an extremely flexible structure of volatility of interest rates. On the downside, many common versions of HJM are non-Markovian, meaning that they are path-dependent, which increases the computational complexity. Also, because the distribution of interest rates is normal, HJM models permit negative interest rates.¹ We shall see what all of this means in this note.

Basic Structure of the Model

In contrast to most other models, which are based on movements in spot interest rates, the HJM model is driven by movements in forward interest rates. We start by

¹That is, given an interest rate change that is normally distributed, the original rate can eventually be driven below zero. In contrast, with proportional rate changes, the rate cannot go below zero.

defining the model as applying over a period of time $t \in [0, T]$.² Let $P(t, T)$ be the price of a zero coupon bond at time t that pays \$1 at time T . Define $f(t, T)$ to be the continuously compounded forward rate observed at time t for an instantaneous transaction to begin at time T . That is, based on the term structure observed at time t , we observe a forward rate for a transaction to start at T and end an instant later. It is well-known that any discounting can be done by the sequence of forward rates instead of spot rates. Hence, a \$1 face value zero coupon bond's price is given as

$$P(t, T) = \exp\left(-\int_t^T f(t, s) ds\right).$$

That is, we can obtain the zero coupon bond's price by successively discounting at the forward rates. Note that we can extract a given forward rate by differentiating the above with respect to T :

$$f(t, T) = -\frac{\partial \log P(t, T)}{\partial T}.$$

While this is a nice and formal definition of a forward rate, it is not of much practical use without a formula that relates the spot price to its maturity. Such a formula can be obtained only for limited cases.³ This formula reminds us, however, that forward rates (specifically, instantaneous forward rates) are (calculus) derivatives of bond prices with respect to maturity.

The shortest forward rate, the one defined as $f(t, t)$ is of special significance and is called the spot rate:

$$r(t) = f(t, t).$$

This is the only spot rate we shall really pay any attention to.

²What we mean here can be expressed in simple terms. If we were just going to examine the term structure, we would pick a series of bonds whose maturities range from one to so many years. Let us call the maturity of the longest maturity bond T . That might not necessarily be the longest maturity bond that exists, but it would be one likely chosen for practical reasons, such as data availability or liquidity. In some cases, we might just need to observe the shortest end of the term structure and T would be short. Regardless, when talking about the term structure, we pick a starting point, which is always time 0, and an ending point, which is the maturity of the longest maturity bond. As we shall see here, to fit the HJM model, we shall need to have bond prices and volatilities that go just past the ending point.

³In most other cases, numerical methods such as binomial models can be used to obtain the price. Partial derivatives can then be estimated by numerical approximations.

Starting from an initial state at time 0, HJM propose that the forward rate observed at time 0 for period T, $f(0,T)$, changes in the following manner during the time from 0 to T:

$$f(t,T) - f(0,T) = \int_0^t \alpha(v,T)dv + \sum_{i=1}^n \int_0^t \sigma_i(v,T)dW_i(v).$$

We now must carefully examine this important equation. Start with the second expression on the right-hand side. Note that it begins with a summation of n terms. This is an n-factor model. These factors are captured by the next terms, $\sigma_i(v,T)$, and $dW_i(v)$. The term $\sigma_i(v,T)$ is the volatility of factor i observed at time v for the forward rate at T.⁴ The term $dW_i(v)$ is a Weiner process representing the source of uncertainty for factor i at time v. There are some formal mathematical restrictions required to uphold these assumptions, but we need not concern ourselves with them here. The expression above in simple terms says that the forward rate started off at a value of $f(0,T)$ and evolved over time to a value of $f(t,T)$. These changes in the forward rate reflected the accumulation, as indicated by the integrals, of the infinitesimal changes that consist of drift and volatility, that have occurred over the period 0 to T.

HJM go on to show that even though we do not really need it, the spot rate process can be derived and is quite similar at

$$r(t) = f(0,t) + \int_0^t \alpha(v,t)dv + \sum_{i=1}^n \int_0^t \sigma_i(v,t)dW_i(v).$$

In addition, the process for the bond price is given as

$$\frac{dP(t,T)}{P(t,T)} = [r(t) + b(t,T)]dt + \sum_{i=1}^n a_i(t,T)dW_i(t).$$

⁴Note that the volatility is not necessarily constant across the term structure or across time. In other words, volatility can change, but it cannot change independent of the level of rates. That is, volatility cannot be independently stochastic. What we mean by this is when volatility is stochastic but unrelated to the level of rates. Some volatility structures have been proposed in which the volatility is functionally related to the level of rates. In that case, the volatility will be stochastic but all of the uncertainty is coming from the uncertainty of the interest rate. Models with stochastically dependent volatility are permitted and can oftentimes be easily accommodated because the uncertainty is not over and above that already present in the model. In the case of non-stochastic volatility, all volatilities are known but they are allowed to change, as long as the change is known. Also, in the version we present here, there is no distinction between states. That is, volatility is the same at a given time regardless of what level rates are at. In the full HJM model, volatility can also differ by states, though it still has to be deterministic or dependently stochastic. As we note in a few sentences, our volatilities will be associated, not with rates, but with factors driving these rates.

where

$$a_i(t, T) \equiv - \int_t^T \sigma_i(t, v) dv$$

$$b(t, T) \equiv - \int_t^T \alpha(t, v) dv + (1/2) \sum_{i=1}^n a_i(t, T)^2.$$

This expression should look somewhat familiar as it resembles the stochastic process ordinarily used for a stock. Note that the drift, however, consists of the risk-free rate and another term, and that there are multiple volatilities representing the multiple factors.

They go on to derive their most important result, which is that the condition of no arbitrage implies that a martingale probability measure exists and implies a restriction on the drift coefficients of the forward rates. Specifically, for the n-factor model

$$\alpha(t, T) = \sum_{i=1}^n \sigma_i(t, T) \int_t^T \sigma_i(t, v) dv.$$

This statements means that the drift is not independent of the volatility and is in fact a specific function of the volatility.⁵ At a given time point t, the drift for the forward rate to start at T is obtained by integrating (essentially adding) all of the volatilities over the time periods from t to T and multiplying by the sum of the volatilities across all of the factors observed at t for the rate to start at T. Again, this restriction assures that no arbitrage opportunities are possible. We do not explore the details of how this is obtained in continuous time, but we will look at it more carefully in a discrete time framework.

In general, the continuous time HJM model is written as

$$df(t, T) = \alpha(t, T)dt + \sum_{i=1}^n \sigma_i(t, T)dW_i(t)$$

with the drift restricted as given above.

In the HJM model the volatility structure is given. What we mean by a volatility structure is a concept with three dimensions. First, there is the element of cross-sectional volatility, referring to the volatility of factor i at a given time point for a set of different forward rates. For example, cross-sectional volatility at time 0 would be indicated by the

⁵For some reason, the literature has stressed the point that the drift under the equivalent martingale measure cannot be zero. HJM (1991) and Ritchken (1996) use a simple binomial tree example to show that if a drift of zero is assumed, there is an arbitrage opportunity. It is not clear why anyone would think that the drift should be zero. From the above formula, it should be clear that the drift cannot be zero if interest rates are stochastic.

variables $\sigma_i(0,1), \sigma_i(0,2), \dots, \sigma_i(0,T)$ for example. These are volatilities associated with forward rates for times 1, 2, ..., T. Cross-sectional volatility would also be indicated at different points in time, such as with $\sigma_i(1,1), \sigma_i(1,2), \dots, \sigma_i(1,T)$. Second, there is the element of times series volatility, referring to the volatility of a factor associated with given forward rate over time. For example, the forward rate maturing at T has a volatility for factor i over time of $\sigma_i(0,T), \sigma_i(1,T), \dots, \sigma_i(T,T)$. Third, if it is indeed a multi-factor version of HJM, there is the concept of factor volatility. If there are n factors, there is a different set of times series and cross-sectional volatilities for each factor i.

There need not be any formal mathematical structure to these volatilities. In other words, they need not be related to each other mathematically. There are some special cases of HJM involving specific mathematical structures to the volatility, such as

$$\sigma(t,T) = \sigma \exp[-\lambda(T-t)],$$

which is called exponentially dampened volatility. In this case a single volatility, σ , is given and all successive volatilities are related to it. The above specification results in volatilities declining at an exponential rate. This particular structure is especially convenient as it permits many closed-form solutions for options and other derivatives. For details see Jarrow and Turnbull (2000, Chs. 16 & 17).

Other volatility functions include the simple case of constant volatility

$$\sigma(t,T) = \sigma,$$

which makes this model equivalent to Ho-Lee. Another case sometimes seen is the nearly proportional volatility,

$$\sigma(t,T) = \eta(t,T) \min(f(t,T), M)$$

where $\eta(t,T)$ is a deterministic function and M is a large positive constant. This specification sets the volatility proportional to the current forward rate and bounds it on the upper end so that it will not get unreasonably high. Other structures seen are

$$\begin{aligned} \sigma(t,T) &= \sigma f(t,T)^\gamma \\ \sigma(t,T) &= \sigma \sigma(t)^\gamma \exp[-\lambda(T-t)]. \end{aligned}$$

Note that these two examples and the nearly proportional volatility case are cases in which the volatility is stochastic but completely dependent on the level of rates. Thus, while these are stochastic volatility models, they are not independent stochastic volatility

models. Hence, they do not pose any additional problems not already present in the model.

For more on volatility structures of the HJM model, see Jarrow (1996), Ritchken (1996) and Richtken and Sankarasubramanian (1995).

Discretizing the HJM Model

With the exception of a few restrictive volatility structures, the HJM model does not produce closed-form solutions for the prices and risk measures of various instruments.⁶ Hence, numerical methods are normally required. Here we will look at discretizing the HJM model for use in a binomial tree. In doing so we shall gain a deeper understanding of the model and especially the drift restriction. At this level, we shall focus exclusively on the one-factor version. Hence, we are given the stochastic process for the forward rate,

$$df(t, T) = \alpha(t, T)dt + \sigma(t, T)dW(t),$$

where the arbitrage-free drift restriction is, again,

$$\alpha(t, T) = \sigma(t, T) \int_t^T \sigma(t, T) dv..$$

To generate a binomial version of the model, we must first consider how much information we have to work with. We shall need the prices of a number of bonds maturing at discrete time points, 1, 2, 3, ..., T-1, T. If T is the longest maturity available, that would mean we have T forward rates available, $f(0,0)$, $f(0,1)$, ..., $f(0,T-1)$.⁷ We would then need volatilities for maturities of 1, 2, ..., T-1. This amount of information would be sufficient to build a binomial tree of T-1 time steps.⁸

First, let us write the stochastic process for the forward rates in discrete form as

$$\Delta f(t, T) = \alpha(t, T)\Delta t + \sigma(t, T)\Delta W(t).$$

⁶As noted, the exponentially decaying volatility structure produces a number of closed-form solutions. See Jarrow and Turnbull (2000, Chs. 16, 17) for details. In addition, Brenner and Jarrow (1993) obtain some closed-form solutions for a special case of a two-factor model. The option pricing formulas in both of these cases bear a striking resemblance to the Black-Scholes model.

⁷If we had T forward rates, this would mean we have T bond prices. Either would imply the other.

⁸Technically, we can go one more step to time T but the only bond that would exist over the final time period is a one-period zero coupon bond, which was the original T-period zero coupon bond at time 0. With one period to go, this bond would have no uncertainty. Thus, for that last period all we would have is the riskless asset. So, all we can really build is a model with T-1 time steps with bonds of maturity up to T.

To binomialize the Wiener process we simply convert it to a random variable with a value of +1 or -1 at each time step and assume martingale probabilities of 1/2.⁹ We assume each time step has a defined length of one unit. Hence, the stochastic process becomes

$$\Delta f(t, T) = \alpha(t, T) \pm \sigma(t, T).$$

Thus, at a given time, for given a forward rate $f(t, T)$, we move one step ahead to the next time in the following manner

$$\begin{aligned} f(t+1, T)^+ &= f(t, T) + \alpha(t, T) + \sigma(t, T) \\ f(t+1, T)^- &= f(t, T) + \alpha(t, T) - \sigma(t, T). \end{aligned}$$

Recall that to prevent arbitrage in a term structure, the local expectations hypothesis (LEH) must hold. It says that the expected return on any financial instrument over the shortest time period must be the riskless rate, where expectations are taken using a special martingale probability measure.¹⁰ That is,

$$P(t, T) = P(t, t+h) E^Q [P(t+h, T)],$$

where the exponent Q means that expectations are taken using the martingale probabilities.¹¹ The first term, $P(t, t+h)$, is the price of the bond with the shortest maturity. By multiplying something by it, we are discounting at the riskless rate. The second term, $E^Q [P(t+h, T)]$ is the expectation of the bond's price at time $t+h$. By discounting this expectation at the riskless rate, we obtain the current bond price.

Writing the above expression as

$$\frac{P(t, T)}{P(t, t+h)} = E^Q [P(t+h, T)],$$

and using the fact that

$$P(t, T) = \exp \left[- \int_t^T f(t, v) dv \right],$$

⁹For a review of why this works, see TN00-05: Brownian Motion: From Discrete to Continuous Time.

¹⁰For more on the local expectations hypothesis, see TN00-02: The Local Expectations Hypothesis. In short, if the LEH holds, the expected return, using the martingale probability measure, over the shortest holding period possible from any bond, bond trading strategy, or combination of bonds is the riskless rate over that holding period. It also means that the forward price one period ahead is the expected spot price, with expectations based on the martingale probability measure. The absence of arbitrage opportunities implies the LEH and the LEH implies the absence of arbitrage opportunities.

¹¹From this point on, we shall always just refer to probabilities and expectations without reference to the fact that they refer to the martingale measure.

we substitute and obtain

$$\exp\left[-\int_t^T f(t, \nu) d\nu\right] \exp\left[\int_t^{t+h} f(t, \nu) d\nu\right] = \exp\left[-\int_{t+h}^T f(t, \nu) d\nu\right].$$

This must equal the expectation given above, which can be found by evaluating the following

$$(1/2) \exp\left[-\int_{t+h}^T (f(t, \nu) + \alpha(t, \nu) + \sigma(t, \nu)) d\nu\right] + (1/2) \exp\left[-\int_{t+h}^T (f(t, \nu) + \alpha(t, \nu) - \sigma(t, \nu)) d\nu\right].$$

This expectation reflects the binomial probabilities of $\frac{1}{2}$ and the integrals represent the discounting of the sequence of forward rates. In other words, the two terms that are multiplied by $\frac{1}{2}$ are the next two possible bond prices, which themselves are obtained by discounting at the sequence of forward rates over the remaining lives of the bonds.

After a fair amount of additional math, we obtain

$$\alpha(t, T) = \sigma(t, T) \int_t^T \sigma(t, \nu) d\nu$$

which is the result we obtained before.¹²

To actually work with the HJM model in discrete time, however, requires that we obtain a discretized version of the drift restriction. Heath, Jarrow, Morton (1991) provide a version of this result, obtained in a clever way. Using a binomial model, they obtain a result that they extend to continuous time by letting the time step approach zero. They then use it to show the correct drift in a simple binomial example. Ritchken (1996, pp. 579-580) repeats this result but both are apparently unaware that this is not correct for the binomial case. Starting with a binomial model, obtaining a result and taking the continuous time limit to that result is certainly correct. But the discrete time version of this continuous time limit is not correct, as shown by Grant and Vora (1999a, 1996b), who go on to derive the correct discrete time formula. They start with a variation of the expression we used above

$$P(t, T) = E^Q[P(t+h, T)]P(t, t+h).$$

¹²See Ritchken (1996, Ch. 24) for details.

They then make use of the fact that a Wiener process is a normal distribution, the interaction of the volatilities has convenient properties, and that the correlation between all forward rates in a one-factor model is 1.0. After considerable algebra, they find a simple expression for the drift

$$\alpha(t, T) = (1/2) \left[\sigma^2(t, T) + 2\sigma(t, T) \sum_{j=t+1}^{T-1} \sigma(t, j) \right].$$

This value can be computed very easily from the covariance matrix of forward rate volatilities, which we shall demonstrate in a numerical example below. Grant and Vora call this the *drift adjustment term* though there is really no reason to call it anything other than the drift.

Fitting a Binomial Tree to the Heath-Jarrow-Morton Model

The following information is given for the term structure.

$$P(0,1) = 0.9343$$

$$P(0,2) = 0.8694$$

$$P(0,3) = 0.8025$$

$$P(0,4) = 0.7393$$

The forward rates are

$$f(0,0) = 0.068 \text{ (which is, of course, the spot rate } r(0))$$

$$f(0,1) = 0.072$$

$$f(0,2) = 0.08$$

$$f(0,3) = 0.082$$

These rates are consistent with the above prices.¹³ The volatilities at time 0 are

$$\sigma(0,1) = 0.02$$

$$\sigma(0,2) = 0.015$$

$$\sigma(0,3) = 0.01$$

The diagram below shows the structure of the problem.

	$f(1,1)^+ = 0.072 + \alpha(0,1) + \sigma(0,1)$ $f(1,2)^+ = 0.08 + \alpha(0,2) + \sigma(0,2)$
--	---

¹³For example, $P(0,4) = \exp(-0.082)\exp(-0.08)\exp(-0.072)\exp(-0.068) = 0.7343$. You may wish to check the remaining prices.

	$f(1,3)^+ = 0.082 + \alpha(0,3) + \sigma(0,3)$
$f(0,0) = 0.068$ $f(0,1) = 0.072$ $f(0,2) = 0.08$ $f(0,3) = 0.082$	
	$f(1,1)^- = 0.072 + \alpha(0,1) - \sigma(0,1)$ $f(1,2)^- = 0.08 + \alpha(0,2) - \sigma(0,2)$ $f(1,3)^- = 0.082 + \alpha(0,3) - \sigma(0,3)$

We need to find the drifts, $\alpha(0,1)$, $\alpha(0,2)$, and $\alpha(0,3)$ to determine the rates at time 1. When we have that done, we can move forward to time 2.

The Grant-Vora formula was given as

$$\alpha(t, T) = (1/2) \left[\sigma^2(t, T) + 2\sigma(t, T) \sum_{j=t+1}^{T-1} \sigma(t, j) \right].$$

Unfortunately, this formula will be a little confusing when $t = 0$ and $T = 1$ because then we are summing from $j = 1$ to 0. What happens is that this whole term drops out. To avoid this confusion, here is an alternative equivalent formula:

$$\alpha(t, T) = \sigma(t, T) \sum_{j=t+1}^T \sigma(t, j) - \sigma^2(t, T) / 2.$$

Moreover, if you are used to working with variances and covariances these formulas may start looking familiar. Consider any two forward rates $f(t, T)$ and $f(t, T-1)$. The covariance at t between these rates is $\sigma(t, T)\sigma(t, T-1)\text{corr}(T, T-1)$ where $\text{corr}(T, T-1)$ is the correlation between the two forward rates. The use of a linear one-factor model, however, means that these rates are perfectly correlated. Hence, the covariance is simply the product of the volatilities, $\sigma(t, T)\sigma(t, T-1)$. Thus, the covariance matrix will be of the form

	1	2	3
1	$\sigma^2(0,1)$	$\sigma(0,1)\sigma(0,2)$	$\sigma(0,1)\sigma(0,3)$
2	$\sigma(0,2)\sigma(0,1)$	$\sigma^2(0,2)$	$\sigma(0,2)\sigma(0,3)$
3	$\sigma(0,3)\sigma(0,1)$	$\sigma(0,3)\sigma(0,2)$	$\sigma^2(0,3)$

In other words, by knowing the volatility structure we know the variances and covariances of all of the forward rates for all time points up to the final one. Grant and Vora show that the drift for any maturity T will be one-half the sum of the elements in the i^{th} row and j^{th} column. For our example, the covariance matrix will be

	1	2	3
1	0.0004	0.0003	0.0002
2	0.0003	0.000225	0.00015
3	0.0002	0.00015	0.0001

Let us begin for maturity $T = 1$. By the formula

$$\begin{aligned}\alpha(0,1) &= \sigma(0,1) \sum_{j=1}^1 \sigma(0,j) - \sigma^2(0,1) / 2 \\ &= \sigma(0,1)\sigma(0,1) - \sigma^2(0,1) / 2 = \sigma^2(0,1) / 2 = (.02)^2 / 2 = 0.0002,\end{aligned}$$

or using the covariance matrix

$$\alpha(0,1) = (1/2)(.0004) = 0.0002.$$

For maturity $T = 2$, by the formula

$$\begin{aligned}\alpha(0,2) &= \sigma(0,2) \sum_{j=1}^2 \sigma(0,j) - \sigma^2(0,2) / 2 \\ &= \sigma(0,2)[\sigma(0,1) + \sigma(0,2)] - \sigma^2(0,2) / 2 \\ &= 0.015(0.02 + 0.015) - 0.000225 / 2 = 0.0004125,\end{aligned}$$

or using the covariance matrix,

$$\alpha(0,2) = (1/2)(0.0003 + 0.000225 + 0.0003) = 0.0004125.$$

For maturity $T = 3$, by the formula

$$\begin{aligned}\alpha(0,3) &= \sigma(0,3) \sum_{j=1}^3 \sigma(0,j) - \sigma^2(0,3) / 2 \\ &= \sigma(0,3)[\sigma(0,1) + \sigma(0,2) + \sigma(0,3)] - \sigma^2(0,3) / 2 \\ &= 0.01(0.02 + 0.015 + 0.01) - 0.0001 / 2 = 0.0004,\end{aligned}$$

or using the covariance matrix,

$$\alpha(0,3) = (1/2)(0.0002 + 0.00015 + 0.0001 + 0.0002 + 0.00015) = 0.0004.$$

So now we can fill in the first time step of the tree.

	$f(1,1)^+ = 0.072 + 0.0002 + 0.02 = 0.0922$ $f(1,2)^+ = 0.08 + 0.0004125 + 0.015 = 0.0954125$ $f(1,3)^+ = 0.082 + 0.0004 + 0.01 = 0.0924$
$f(0,0) = 0.068$ $f(0,1) = 0.072$ $f(0,2) = 0.08$ $f(0,3) = 0.082$	
	$f(1,1)^- = 0.072 + 0.0002 - 0.02 = 0.0522$ $f(1,2)^- = 0.08 + 0.0004125 - 0.015 = 0.0654125$ $f(1,3)^- = 0.082 + 0.0004 - 0.01 = 0.0724$

Now suppose we wanted to move forward another time step. If we position ourselves at the top step at time 1 (where the rates are $f(1,1)^+$, $f(1,2)^+$, and $f(1,3)^+$) we can then imagine moving forward to time step 2 in the following manner.

	$f(2,2)^{++} = 0.0954125 + \alpha(1,2) + \sigma(1,2)$ $f(2,3)^{++} = 0.0924 + \alpha(1,2) + \sigma(1,2)$
$f(1,1)^+ = 0.0922$ $f(1,2)^+ = 0.0954125$ $f(1,3)^+ = 0.0924$	
	$f(2,2)^{+-} = 0.0954125 + \alpha(1,2) - \sigma(1,2)$ $f(2,3)^{+-} = 0.0924 + \alpha(1,3) - \sigma(1,3)$

Note that we now require not only new drift terms, but we see volatility terms not previously seen. That is, we started with a term structure of volatility of $\sigma(0,1)$, $\sigma(0,2)$, and $\sigma(0,3)$. These were the volatilities from time point 0. Now, having moved forward to time point 1, we have a new set of volatilities $\sigma(1,2)$ and $\sigma(1,3)$. Recall that $\sigma(0,2)$ was our value at time 0 of the volatility of the forward rate for time 2. Now, at time 1, $\sigma(1,2)$ is our value of that volatility. Should it be different? In HJM it can indeed be different. Volatility can change over time, but it cannot change stochastically. That is, while $\sigma(1,2)$

can be different from $\sigma(0,2)$, we had to have known both values at time 0. In other words, volatility can change, but we have to know what it will change to.¹⁴

If we do not impose any changes on volatility then we simply use the value $\sigma(0,2)$ for $\sigma(1,2)$ and use $\sigma(0,3)$ for $\sigma(1,3)$. Then we would be able to calculate the drift and fit the tree. Then we would step down to the lower state and time 1 and determine the rates in the next two states at time 2. As long as we make the assumption that $\sigma(0,2) = \sigma(1,2)$ and $\sigma(0,3) = \sigma(1,3)$, the tree will recombine.¹⁵ In other words, we require constant time series volatility. We do not, however, require equivalent cross-sectional volatility. In other words, all rates do not have to have the same volatility for the tree to recombine, but a given rate must have the same volatility across time for the tree to recombine.¹⁶ Also, note that as we move to time step 3, we would require $\sigma(2,3)$, which might also be assumed equal to $\sigma(0,3)$ or could be treated as a different value.

Once we have the entire tree filled in, we have all of the information we need to price any bond or derivative. The rest of the problem is illustrated in the appendix. Recall that we never really worked with spot rates. They are related to the forward rates, but all instruments can be priced off of the forward rates as easily as off of the spot rates. In other words, there is no more information in the spot rates that we do not already have with the forward rates.

We should also note that because of the assumption of a normal distribution of interest rates, it is possible to obtain negative interest rates. This problem, as well as the problem of the tree not recombining, can be addressed in certain versions of the model.¹⁷

Here we have worked only with one-factor models. In the language of terms structure movements, single factor models seem to capture only changes in the general level of interest rates, but do not reflect changes in the slope or curvature of the term

¹⁴Recall that what we mean is that volatility cannot change stochastically and independently of the level of rates.

¹⁵For example, consider the case of $f(2,2)^+$ vs. $f(2,2)^-$. The former is $f(1,2)^+ + \alpha(1,2) - \sigma(1,2)$. Substituting $f(0,2) + \alpha(0,2) + \sigma(0,2)$ for $f(1,2)^+$ means that $f(1,2)^+ = f(0,2) + \alpha(0,2) + \sigma(0,2) + \alpha(1,2) - \sigma(1,2)$. A similar operation reveals that $f(1,2)^- = f(0,2) + \alpha(0,2) - \sigma(0,2) + \alpha(1,2) + \sigma(1,2)$. These two expressions are equivalent if and only if $\sigma(0,2) = \sigma(1,2)$. A similar result is found for $f(2,3)^+$ vs. $f(2,3)^-$, requiring that $\sigma(0,3) = \sigma(1,3)$.

¹⁶In general, for T time steps, a non-recombining tree would have 2^T paths.

¹⁷See de Munnik (1994) for a recombining tree. The problem of negative rates can generally be addressed by volatility specifications that dampen the volatility sufficiently to prevent it from moving further downward when at or near zero.

structure. To capture these effects, multi-factor models are required. Multi-factor models are much more complex. For example, a tree version of a two-factor HJM model requires a trinomial, and the number of paths for T time steps is 3^T . Clearly some tradeoffs are required before deciding to go to models of more than one factor.

References

The original work on the model is the three classic papers by HJM:

Heath, D., R. A. Jarrow, and A. Morton. "Contingent Claims Valuation with a Random Evolution of Interest Rates." *Review of Futures Markets* 9 (1991), 54-76.

Heath, D., R. Jarrow, and A. Morton. "Bond Pricing and the Term Structure of Interest Rates: A New Methodology." *Econometrica* 60 (1992), 77-105.

Heath, D., R. Jarrow, and A. Morton. "Bond Pricing and the Term Structure of Interest Rates: A Discrete Time Approximation." *Journal of Financial and Quantitative Analysis* 25 (1990), 419-440.

Other useful work is

Grant, D. and G. Vora. "Extending the Universality of the Heath-Jarrow-Morton Model." Working paper, University of New Mexico (1999a).

Grant, D. and G. Vora. "Implementing No-Arbitrage Terms Structure of Interest Rate Models in Discrete Time When Interest Rates are Normally Distributed." *The Journal of Fixed Income* 8 (May, 1999b), 85-98.

Ritchken, P. *Derivative Markets*. New York: HarperCollins (1996), Ch. 24.

Jarrow, R. *Modelling Fixed Income Securities and Interest Rate Options*. New York: McGraw-Hill (1996), Ch. 12.

Jarrow, R. and S. Turnbull. *Derivative Securities*. Cincinnati: South-Western College Publishing (2000), Ch. 16.

Some advanced papers are

Brenner, R. J. and R. A. Jarrow. "A Simple Formula for Options on Discount Bonds." *Advances in Futures and Options Research* 6 (1993), 45-51.

Carverhill, A. "A Simplified Exposition of the Heath, Jarrow, and Morton Model." *Stochastics and Stochastics Reports* 53 (1995), 227-240.

de Munnik, J. F. J. “The Construction of a Path-Independent Interest Rate Tree: The Model of Heath, Jarrow, and Morton.” *Advances in Futures and Options Research* 7 (1994), 135-145.

Ritchken, P. and L. Sankarasubramanian. “Volatility Structures of Forward Rates and the Dynamics of the Term Structure.” *Mathematical Finance* 5 (1995), 55-72.

HJM explain how to use their model in relatively simple terms in

Heath, D., R. Jarrow, A. Morton, and M. Spindel. “Easier Done Than Said.” *Risk* 5 (October, 1992), 77-80.

There is much more written about the model and most advanced books on derivatives cover it quite thoroughly.

Appendix: The Remainder of the HJM Binomial Tree

Recall that in the text of this note, we obtained the following rates:

$$\begin{aligned} f(1,1)^+ &= f(0,1) + \alpha(0,1) + \sigma(0,1) = 0.072 + 0.0002 + 0.02 = 0.0922 \\ f(1,2)^+ &= f(0,2) + \alpha(0,2) + \sigma(0,2) = 0.08 + 0.0004125 + 0.015 = 0.0954125 \\ f(1,3)^+ &= f(0,3) + \alpha(0,3) + \sigma(0,3) = 0.082 + 0.0004 + 0.01 = 0.0924 \end{aligned}$$

$$f(0,0) = 0.068$$

$$f(0,1) = 0.072$$

$$f(0,2) = 0.08$$

$$f(0,3) = 0.082$$

$$f(1,1)^- = f(0,1) + \alpha(0,1) - \sigma(0,1) = 0.072 + 0.0002 - 0.02 = 0.0522$$

$$f(1,2)^- = f(0,1) + \alpha(0,1) - \sigma(0,2) = 0.08 + 0.0004125 - 0.015 = 0.0654125$$

$$f(1,3)^- = f(0,1) + \alpha(0,1) - \sigma(0,3) = 0.082 + 0.0004 - 0.01 = 0.0724$$

We will be adding the following:

$$\begin{aligned} f(2,2)^{++} &= f(1,2)^+ + \alpha(1,2) + \sigma(1,2) \\ f(2,3)^{++} &= f(1,3)^+ + \alpha(1,3) + \sigma(1,3) \\ f(2,2)^{+-} &= f(1,2)^+ + \alpha(1,2) - \sigma(1,2) \\ f(2,3)^{+-} &= f(1,3)^+ + \alpha(1,3) - \sigma(1,3) \\ f(2,2)^{-+} &= f(1,2)^- + \alpha(1,2) + \sigma(1,2) \\ f(2,3)^{-+} &= f(1,3)^- + \alpha(1,3) + \sigma(1,3) \\ f(2,2)^-- &= f(1,2)^- + \alpha(1,2) - \sigma(1,2) \\ f(2,3)^-- &= f(1,3)^- + \alpha(1,3) - \sigma(1,3) \\ f(3,3)^{+++} &= f(2,3)^{++} + \alpha(2,3) + \sigma(2,3) \\ f(3,3)^{++-} &= f(2,3)^{++} + \alpha(2,3) - \sigma(2,3) \\ f(3,3)^{-++} &= f(2,3)^{-+} + \alpha(2,3) + \sigma(2,3) \\ f(3,3)^{-+-} &= f(2,3)^{-+} + \alpha(2,3) - \sigma(2,3) \\ f(3,3)^{+--} &= f(2,3)^+ + \alpha(2,3) + \sigma(2,3) \\ f(3,3)^{-+-} &= f(2,3)^- + \alpha(2,3) - \sigma(2,3) \end{aligned}$$

$$\begin{aligned}f(3,3)^{-+} &= f(2,3)^{-} + \alpha(2,3) + \sigma(2,3) \\f(3,3)^{-} &= f(2,3)^{-} + \alpha(2,3) - \sigma(2,3)\end{aligned}$$

In order to get the drifts at times 1 and 2, we need the volatilities $\sigma(1,2)$, $\sigma(1,3)$, and $\sigma(2,3)$. Making the assumption that the volatilities do not change, then

$$\begin{aligned}\sigma(0,2) &= \sigma(1,2) = \sigma(2,2) \\ \sigma(0,3) &= \sigma(1,3) = \sigma(2,3)\end{aligned}$$

The covariance matrix of rates at time 1 looks as follows:

$$\begin{array}{cc} \sigma^2(1,2) & \sigma(1,2)\sigma(1,3) \\ \sigma(1,3)\sigma(1,2) & \sigma^2(1,3) \end{array}$$

The term $\sigma^2(1,2)$ is the variance of the one-period rate at time 1, and $\sigma^2(1,3)$ is the variance of the two-period rate at time 1. The term $\sigma(1,2)\sigma(1,3)$, which equals $\sigma(1,3)\sigma(1,2)$, is the covariance of the one- and two-period rates at time 1. Filling in the numbers we have

$$\begin{aligned}\sigma^2(1,2) &= 0.000225 & \sigma(1,2)\sigma(1,3) &= (0.015)(0.01) = 0.00015 \\ \sigma(1,3)\sigma(1,2) &= (0.01)(0.015) = 0.00015 & \sigma^2(1,3) &= 0.0001\end{aligned}$$

Recall that in the text we explained that the drift can be obtained by taking $\frac{1}{2}$ of the sum of the terms in the appropriate row and column. That is, if we want the drift $\alpha(i,j)$, we take $\frac{1}{2}$ the sum of the elements in the i^{th} row and j^{th} column, counting element ij only once. Thus, the drifts are obtained as

$$\begin{aligned}\alpha(1,2) &= 0.000225/2 = 0.0001125 \\ \alpha(1,3) &= (1/2)(0.00015 + 0.0001 + 0.00015) = 0.0002\end{aligned}$$

Then the rates at time 2 are

$$\begin{aligned}f(2,2)^{++} &= f(1,2)^+ + \alpha(1,2) + \sigma(1,2) = 0.0954125 + 0.0001125 + 0.015 = 0.110525 \\ f(2,3)^{++} &= f(1,3)^+ + \alpha(1,3) + \sigma(1,3) = 0.0924 + 0.0002 + 0.01 = 0.1026 \\ f(2,2)^+- &= f(1,2)^+ + \alpha(1,2) - \sigma(1,2) = 0.0954125 + 0.0001125 - 0.015 = 0.080525 \\ f(2,3)^+- &= f(1,3)^+ + \alpha(1,3) - \sigma(1,3) = 0.0924 + 0.0002 - 0.01 = 0.0826 \\ f(2,2)^-+ &= f(1,2)^- + \alpha(1,2) + \sigma(1,2) = 0.0654125 + 0.0001125 + 0.015 = 0.080525 \\ f(2,3)^-+ &= f(1,3)^- + \alpha(1,3) + \sigma(1,3) = 0.0724 + 0.0002 + 0.01 = 0.0826 \\ f(2,2)^-- &= f(1,2)^- + \alpha(1,2) - \sigma(1,2) = 0.0654125 + 0.0001125 - 0.015 = 0.050525 \\ f(2,3)^-- &= f(1,3)^- + \alpha(1,3) - \sigma(1,3) = 0.0724 + 0.0002 - 0.01 = 0.0626\end{aligned}$$

Note that because of the constant volatility assumption, some of these rates are duplicates. That is, $f(2,2)^{-+} = f(2,2)^{-}$ and $f(2,3)^{-+} = f(2,3)^{-}$.

Moving on to time 3, the covariance matrix of rates at time 2 is simple. We have only the rate $\sigma^2(2,3)$, which is 0.0001. Then the drift, $\alpha(2,3) = 0.0001/2 = 0.00005$. The rates will be

$$\begin{aligned} f(3,3)^{+++} &= f(2,3)^{++} + \alpha(2,3) + \sigma(2,3) = 0.1026 + 0.00005 + 0.01 = 0.11265 \\ f(3,3)^{++-} &= f(2,3)^{++} + \alpha(2,3) - \sigma(2,3) = 0.1026 + 0.00005 - 0.01 = 0.09265 \\ f(3,3)^{+-+} &= f(2,3)^{+-} + \alpha(2,3) + \sigma(2,3) = 0.0826 + 0.00005 + 0.01 = 0.09265 \\ f(3,3)^{-+-} &= f(2,3)^{+-} + \alpha(2,3) - \sigma(2,3) = 0.0826 + 0.00005 - 0.01 = 0.07265 \\ f(3,3)^{+--} &= f(2,3)^{+-} + \alpha(2,3) + \sigma(2,3) = 0.0826 + 0.00005 + 0.01 = 0.09265 \\ f(3,3)^{-+-} &= f(2,3)^{-+} + \alpha(2,3) - \sigma(2,3) = 0.0826 + 0.00005 - 0.01 = 0.07265 \\ f(3,3)^{-+} &= f(2,3)^{-+} + \alpha(2,3) + \sigma(2,3) = 0.0626 + 0.00005 + 0.01 = 0.07265 \\ f(3,3)^{--} &= f(2,3)^{-} + \alpha(2,3) - \sigma(2,3) = 0.0626 + 0.00005 - 0.01 = 0.05265 \end{aligned}$$

Note here that the recombining tree means that $f(3,3)^{++-} = f(3,3)^{+-+} = f(3,3)^{-+-}$, $f(3,3)^{-+-} = f(3,3)^{-+} = f(3,3)^{-+}$.

The tree looks as follows:

$$\begin{array}{rcl} & & f(3,3)^{+++} = 0.1126 \\ & & f(2,2)^{++} = 0.110525 \\ & & f(2,3)^{++} = 0.1026 \\ & f(1,1)^+ = 0.0922 & \\ & f(1,2)^+ = 0.0954125 & f(3,3)^{++-} = 0.09265 \\ & f(1,3)^+ = 0.0924 & \\ f(0,0) = 0.068 & & \\ f(0,1) = 0.072 & f(2,2)^{+-} = 0.080525 & \\ f(0,2) = 0.08 & f(2,3)^{+-} = 0.0826 & \\ f(0,3) = 0.082 & & \\ & f(1,1)^- = 0.0522 & \\ & f(1,2)^- = 0.0654125 & f(3,3)^{-+-} = 0.07265 \\ & f(1,3)^- = 0.0724 & \\ & & f(2,2)^{-} = 0.050525 \\ & & f(2,3)^{-} = 0.0626 \\ & & f(3,3)^{-+} = 0.07265 \\ & & f(2,2)^{-} = 0.050525 \\ & & f(2,3)^{-} = 0.0626 \\ & & f(3,3)^{-} = 0.05265 \end{array}$$

Now we have fit the entire term structure. If we have done this correctly, it should be arbitrage-free. If that is the case, then we should be able to derive a tree for the price of any instrument, such that each price is the discounted value of the expected price the next period, where expectations are taken using the martingale probability of $\frac{1}{2}$. We shall do this for a four-period zero coupon bond. First, let us establish its price at time 0. The initial information given is the set of forward rates $f(0,0)$, $f(0,1)$, $f(0,2)$, and $f(0,3)$. They imply that the price of the bond at time zero should be

$$P(0,4) = \exp(-0.08 + 0.072 + 0.08 + 0.082) = 0.7393.$$

We shall start by pricing this bond at time 3, then work backwards. When we get to time zero, we should get this price, subject to perhaps a round-off error.

The price at time 4 is obtained by discounting the \$1 face value at the spot rate $f(3,3)$ at time 3. There are, of course, four states at time 3. Thus, the prices are

$$\begin{aligned} P(3,4)^{+++} &= \exp(-f(3,3)^{+++}) = \exp(-0.11265) = 0.8935 \\ P(3,4)^{++-} &= \exp(-f(3,3)^{++-}) = \exp(-0.09265) = 0.9115 \\ P(3,4)^{+-+} &= \exp(-f(3,3)^{+-+}) = \exp(-0.07265) = 0.9300 \\ P(3,4)^{---} &= \exp(-f(3,3)^{---}) = \exp(-0.05265) = 0.9487 \end{aligned}$$

The prices of this bond at time 2 are found by weighting each of the next two possible outcomes by .5 and discounting by the one-period spot rate at time 2, $f(2,2)$:

$$\begin{aligned} P(2,4)^{++} &= (0.5(0.8935) + 0.5(0.9115))\exp(-0.110525) = 0.8081 \\ P(2,4)^{+-} &= (0.5(0.9115) + 0.5(0.9300))\exp(-0.080525) = 0.8495 \\ P(2,4)^{-} &= (0.5(0.9300) + 0.5(0.9487))\exp(-0.050525) = 0.8931 \end{aligned}$$

Stepping back to time 1, we weight each of the next two values by .5 and discount by the one-period spot rate at time 1, $f(1,1)$:

$$\begin{aligned} P(1,4)^{+} &= (0.5(0.8081) + 0.5(0.8495))\exp(-0.0922) = 0.7558 \\ P(1,4)^{-} &= (0.5(0.8495) + 0.5(0.8931))\exp(-0.0522) = 0.8270 \end{aligned}$$

The price at time 0 is found by weighting each of the next two values by .5 and discounting by the one-period spot rate at time 0, $f(0,0)$:

$$P(0,4) = (0.5(0.7558) + 0.5(0.8270))\exp(-0.068) = 0.7394$$

This value differs from the value we originally computed by a round-off error.