Backus, Wu, & Zin

Technical Note on Hull and White

(February 6, 1999; revised April 13, 1999)

Hull and White have translated the Vasicek model (and others) into a discrete-time, discrete-state setting that's easy to program and widely used by professionals. But we doubt we're alone in finding some of the details of their approach a little mysterious. than perfectly transparent. These notes are an attempt to fill in the some of the missing details. For the most part they are self-contained, but we refer to the original figures to save effort.

Summary. The trick is to take a "trinomial tree" and get it to reproduce the main features of a stochastic process with mean reversion. To do this, they play with the probabilities in the tree, raising the probability of a decline (fall) when the state variable z is high (low). Eventually, this drives one of the probabilities negative, so they truncate the tree. This has another benefit: the tree doesn't grow ridiculously large as the number of periods grows. In standard binomial trees, the number of states is quadratic in the number of periods since in period n there are n nodes. Here it is linear in the number of periods because the number of nodes per period has an upper bound. The key assumption is stationarity/mean-reversion: the probability of extreme events doesn't increase without bound in the number of periods, so we can safely bound the tree.

1 A Continuous Process

The starting point is a stochastic process for a continuous state variable z:

$$z_{t+h} - z_t \;=\; -\kappa z_t h + \sigma \varepsilon_{t,t+h}, \tag{1}$$

where h is a (presumably short) time interval measured in annual units and $\{\varepsilon_{t,t+h}\} \sim \text{NID}(0, h)$. Parameter $1 > \kappa > 0$ governs the rate of mean reversion, with large values indicating faster mean reversion. Later on, we let the short rate be z plus a time-dependent constant. Other examples define z as the log of the short rate or other monotonic function.

For future reference, consider the salient properties of the process for z. Over the interval h, the first two (conditional) moments of z are

$$\begin{aligned} E_t\,(z_{t+h} - z_t) &= -\kappa h z_t \;=\; M z_t \\ Var_t\,(z_{t+h} - z_t) &= \sigma^2 h \;=\; V, \end{aligned}$$

which defines $M = -\kappa h < 0$ and $V = \sigma^2 h$. Generally both h and κ are fractions. In any case, we assume $-1 < M < 0$.

2 Mechanics of Trinomial Trees

We're going to mimic this process with a tree. At each node, z can go up ("u"), down ("d"), or stay the same ("m" for middle) with probabilities (π_u, π_m, π_d). The size of the up and down changes in z is δ. We will come back to this later, but for now consider δ as given. A node is identified by (i, n), where n is the number of periods after the start of the tree (ie, we start at zero) and i is the net number of up moves (the number of up moves minus the number of down moves). For a specific n, the tree is constructed so that $|i| \leq n$. For large enough n, there is an upper bound $i_{\max}$ such that $|j| \leq i_{\max}$. If we start z at zero, then at (j, n), $z = j\delta$.

Probabilities

Overview. The trick is to pick probabilities for the tree that reproduce the mean and variance of (1). Two issues arise. (i) We have three parameters (two probabilities and δ) to reproduce two moments (the mean and variance of Δz). Where does the extra parameter come from? (ii) How do we know the probabilities are positive? This puts limits on i.

Interior nodes. At each node, we choose the probabilities to match the mean and variance of changes in z. For a node with arbitrary j (n doesn't matter here, since z is stationary), the mean and variance are

$$\begin{aligned} E(\Delta z) &= \pi_u(i)\delta + \pi_m(i)0 + \pi_d(i)(-\delta) \;=\; M(i\delta) \\ Var(\Delta z) &= E\left[(\Delta z)^2\right] - [E(\Delta z)]^2 \\ &= \pi_u(i)\delta^2 + \pi_m(i)0^2 + \pi_d(i)(-\delta)^2 - (jM\delta)^2 \;=\; V \end{aligned}$$

These equations are linear in (π_u, π_d). At $i = 0$, $\pi_u(0) = \pi_d(0) = V/2\delta^2 = \pi_0$ (say). For arbitrary i, the probabilities are

$$\begin{aligned} \pi_u(i) &= \pi_0 + \frac{1}{2}\left[iM + (iM)^2\right] \\ \pi_m(i) &= (1 - 2\pi_0) - (iM)^2 \\ \pi_d(i) &= \pi_0 + \frac{1}{2}\left[-iM + (iM)^2\right]. \end{aligned} \qquad (2)$$

This version illustrates one of the choices involved in this modeling effort: we can choose the initial up and down probabilities π_0, or we can choose the distance δ between nodes, but we can't choose them independently.

The missing moment. Hull and White set $\pi_0 = 1/6$; the question is why. Our answer (perhaps theirs, too) is that we are trying to approximate a normal process. In addition to mean and variance, we might consider central moments

$$\mu_k \;=\; E\left[\Delta z - E(\Delta z)\right]^k$$

for $k > 2$. With $k = 3$ we are measuring skewness, with $k = 4$ kurtosis. Skewness is zero for both the continuous normal process and trinomial tree: both are symmetric. The choice $\pi_0 = 1/6$ reproduces the kurtosis of the normal ($\mu_4 = 3V^2$) in state $i = 0$ (the initial node). The fourth moment satisfies

$$\mu_4 \;=\; \pi_0 \delta^4 + (1 - 2\pi_0) 0^4 + \pi_0 (-\delta)^4 \;=\; 3V^2.$$

Since $V = \pi_0 2 \delta^2$, we find $\pi_0 = 1/6$ and $\delta = (V/2\pi_0)^{1/2} = (3V)^{1/2}$, as suggested by Hull and White. In states $i \neq 0$, skewness is nonzero and kurtosis is different from the normal. What we care about, however, is the conditional distribution over several periods at the initial node, something we won't go into here. Roughly speaking, however, the central limit theorem makes (average) changes more normal the more periods we have. [Comment: not quite true here because of mean reversion...]

The edges. Hull and White bound the tree at $|i| = i_{\max}$, a fixed number whose value we'll go into later. For now, the question is how we set the probabilities at the edges of the tree. (A picture is useful here: see Exhibit 2 of Hull and White 1994 or Exhibit 1 of Hull and White 1996.) The idea is to point the branches back into the tree: at the upper edge, we have branches pointing (same, down, more down). If we call these by the same names (up, middle, down), the probabilities solve these equations,

$$\begin{aligned} E\,(\Delta z) &= \pi_m(i)(-\delta) + \pi_d(i)(-2\delta) \;=\; M(i\delta) \\ Var\,(\Delta z) &= \pi_m(i)(-\delta)^2 + \pi_d(i)(-2\delta)^2 - (iM\delta)^2 \;=\; V. \end{aligned}$$

The solution is

$$\begin{aligned} \pi_u(i) &= 1 + \pi_0 + \frac{1}{2}\left[3iM + (iM)^2\right] \\ \pi_m(i) &= -2\pi_0 - 2iM - (iM)^2 \\ \pi_d(i) &= \pi_0 + \frac{1}{2}\left[iM + (iM)^2\right]. \end{aligned}$$

At the bottom edge, we reverse π_u and π_d. Note, in general, that $\pi_u(i) = \pi_d(-i)$, so it's sufficient to specify the upper edge.

Finding the Edge

This works as long as the probabilities are all between zero and one, which puts bounds on $|i|$ that limit the size of the tree. The binding constraint is at zero: if a probability hits one, then at least one of the other probabilities must be at or below its lower bound of zero.

Basic logic (for positive i). Starting at i at zero and increasing, the interior probabilities behave as follows:

- π_u is initially decreasing in i, but for $\pi_0 \geq 1/8$ it starts increasing again without going negative.
- π_m is decreasing throughout, hitting zero first at $iM = -(1-2\pi_0)^{1/2}$. With $\pi_0 = 1/6$, i is bounded by
$$i \leq \frac{(2/3)^{1/2}}{-M}. \tag{3}$$
(Remember: M is negative and small.)
- π_d is increasing in i so we can ignore it as a relevant bound on i.

Now consider the edge probabilities:

- π_u initially declines with i, then increases. As long as $\pi_0 \geq 1/8$ it's always positive.
- π_m initially increases, then decreases. For $\pi_0 \geq 1/8$, it takes positive values in the interval
$$\frac{1-(1-2\pi_0)^{1/2}}{-M} \leq i \leq \frac{1+(1-2\pi_0)^{1/2}}{-M} \tag{4}$$
- π_d is positive for all $i \geq 0$ if $\pi_0 \geq 1/8$.

In both the interior and on the edge, then, the binding constraints on i come from the middle probability. The edge places both upper and lower bounds on $i_{\max}$, since for i small, the structure of the tree on the edge (no true up move) is inconsistent with mean reversion (which is small for small i). The interior places an upper bound on $i_{\max} - 1$. The issue is how the two interact. Define $a = (1-2\pi_0)^{1/2}$. Consider the upper bound in relation (4), $i_{\max} = -(1+a)/M$. With this value, there is no guarantee that $i_{\max} - 1$ satisfies the interior condition (3). Alternatively, consider the lower bound in (4), $i_{\max} = -(1-a)/M$. In this case, $i_{\max} - 1$ automatically satisfies (3) as long as $\pi_0 \leq 3/8$ (this is a sufficient condition). We should also worry about i being an integer, but you get the idea.

Perhaps for this reason, Hull and White (1994, p 12) suggest choosing
$$i_{\max} = \frac{1-(1-2\pi_0)^{1/2}}{-M} = \frac{-0.18350}{-M},$$
the last equality stemming from the choice $\pi_0 = 1/6$.

Summary

We take as given values of M and V summarizing, respectively, the mean reversion and volatility of the process for z. Given these values, we set $\pi_0 = 1/6$ and $\delta = (3V)^{1/2}$. The number of states, $i_{\max}$, is set equal to the smallest integer greater than $[(1-2\pi_0)^{1/2}-1]/M = -0.18350/M$. Probabilities are then set by (2) in the interior ($|i| < i_{\max}$) and by (3) on the upper edge ($i = i_{\max}$). On the lower edge, we use the relation $\pi_u(i) = \pi_d(-i)$.

3 Short Rate Trees

We now have a procedure for generating a trinomial tree for the state variable z, by which we mean

- a set of states (i, n),
- numbers $z(i, n) = i\delta$ specifying z's value in state (i, n), and
- probabilities $(\pi_u(i), \pi_m(i), \pi_d(i))$ for all permissible i.

Neither z nor π depends on n, a property referred to as stationarity: the behavior of z depends on its current value (via i), but not on the date.

Given z, we set short rates as follows. Define $z'(i, n) = z(i, n) + \mu_n$, where $\mu = (\mu_0, \mu_1, \ldots)$ is a vector of constants that depend on n but not i. Now define the short rate by one of the following:

$$r(i, n) \ = \ \begin{cases} z'(i, n) \\ \exp[z'(i, n)] \end{cases}$$

The former applies to a linear "Ho and Lee" kind of tree, the latter to a log-linear "Black-Derman-Toy" kind of structure.

Finally (and this is simply a collection of arbitrary conventions), define the one-period discount factor in each state by

$$b(i, n) \ = \ e^{-r(i, n)h/100}.$$

The conventions are that short rates are continuously compounded and reported as annual percentages.

Calibration follows the usual routine: we choose μ to reproduce observed spot rates. We'll do an example shortly.

4 Valuation

Consider a claim to state-contingent cash flows $c(j, n)$. There are two standard approaches to valuation. One is recursive: one period at a time, starting with the last cash flow(s). The other is to apply state prices $Q(i, n)$:

$$p \ = \ \sum_{i,n} Q(i, n)c(i, n) \tag{5}$$

We describe each in turn. We follow (with no particular justification) standard practice in equating true probabilities (above) and risk-neutral probabilities.

With both of these methods, it's convenient to imbed the model in the more general setting of a *Markov chain*. In a Markov chain, a state variable like our z assumes a finite number of values, indexed by i. In the Hull and White model, i is an integer between $-i_{\max}$ and $i_{\max}$. Transition probabilities $\pi(i,j)$ describe the likelihood of moving from a given state i this period to another state j the next:

$$\pi(i,j) \;=\; \text{Prob}(z_{n+1} = z_j | z_n = z_i).$$

The collection of such probabilities can be put into a "transition matrix" Π. In our case, the transition matrix for $i_{\max} = 3$ is

	Future State j						
Current State i	–3	–2	–1	0	1	2	3
–3	$\pi_d(-3)$	$\pi_m(-3)$	$\pi_u(-3)$	0	0	0	0
–2	$\pi_d(-2)$	$\pi_m(-2)$	$\pi_u(-2)$	0	0	0	0
–1	0	$\pi_d(-1)$	$\pi_m(-1)$	$\pi_u(-1)$	0	0	0
0	0	0	$\pi_d(0)$	$\pi_m(0)$	$\pi_u(0)$	0	0
1	0	0	0	$\pi_d(1)$	$\pi_m(1)$	$\pi_u(1)$	0
2	0	0	0	0	$\pi_d(2)$	$\pi_m(2)$	$\pi_u(2)$
3	0	0	0	0	$\pi_d(3)$	$\pi_m(3)$	$\pi_u(3)$

In this way, the probabilities described earlier define the elements $\pi(i,j)$ of the matrix Π. This more general structure saves us work later on in allowing us to ignore things like the fact that state -1 next period can be reached from 4 different states the previous period, but state 0 can only be reached from 3 (read down the appropriate columns).

Recursive valuation

One approach to valuation is to start at the end, and value the cash flows at each node by discounting the cash flows of the nodes stemming from it. In the final period $n_{\max}$, the value of an asset is the value of the cash flow in that state:

$$p(i, n_{\max}) \;=\; c(i, n_{\max}).$$

In each node (i,n) for $n < n_{\max}$, the value is the current cash flow $c(i,n)$ plus the sum of the values of the nodes j leading out of it:

$$p(i,n) \;=\; c(i,n) + b(i,n) \sum_j \pi(i,j) p(j, n+1). \tag{6}$$

(This is where we assume that the π's are the risk-neutral probabilities.) For a typical interior node, we could also write this as

$$p(i,n) \;=\; c(i,n) + b(i,n)\left[\pi_d(i)p(i-1,n+1) + \pi_m(i)p(i,n+1) + \pi_u(i)p(i+1,n+1)\right].$$

The Markov chain notation saves us from having to treat nodes differently depending on how many nodes lead to it: the zeros in Π take care of this automatically.

State Prices

A second approach to valuation is to apply state prices directly to cash flows. Define the state price $Q(i,n)$ as the price in state (0,0) of one dollar payable in state (i,n). State prices are readily computed by

$$Q(j,n+1) \;=\; \sum_i \pi(i,j)b(i,j)Q(i,n) \tag{7}$$

We start at the beginning, with $Q(0,0) = 1$ and $Q(i,0) = 0$ for $i \neq 0$, and use (7) to compute Q one period at a time. The sum over i here means over all states i at date n that can lead to j at $n+1$.

State prices can be applied directly to cash flows to value assets using (5). They can also be used to compute discount factors and spot rates when the model is calibrated. Discount factors and (continuously compounded) spot rates at date $n=0$ are

$$\begin{aligned} b_n &= \sum_j Q(j,n) \\ b_{n+1} &= \sum_j Q(j,n)b(j,n) \\ y_n &= -(100/nh)\log b_n. \end{aligned}$$

We return to this shortly.

5 Calibration: Hull and White's Example

Hull and White (1996) work through an example based on an annual time interval $h=1$. They set $\kappa = 0.1$ and $\sigma = 0.01$ and set

$$\begin{aligned} M &= \exp(-\kappa h) - 1 \;=\; -0.095162 \\ V &= \sigma^2[1-\exp(-2\kappa h)]/2\kappa \;=\; 0.0095202^2 \end{aligned}$$

For small h these expressions are pretty much the same as those given earlier, but we stick with theirs to keep the numbers the same.

Interest rate grid. The distance between interest rates in the tree is $\delta = (3V)^{1/2} = 0.0164895$. The maximum state is the largest integer greater than $0.18350/0.095162 = 1.93$, namely 2. That gives us 5 states, ranging from -2 to 2.

Probabilities. We apply (2,3) to get the transition probabilities

0.8993	0.0111	0.0896	0	0
0.1236	0.6576	0.2188	0	0
0	0.1667	0.6667	0.1667	0
0	0	0.2188	0.6576	0.1236
0	0	0.0896	0.0111	0.8993

Short rate tree. The z tree is

$n = 0$	$n = 1$	$n = 2$	$n = 3$
		3.2979	3.2979
	1.6490	1.6490	1.6490
0	0	0	0
	−1.6490	−1.6490	−1.6490
		−3.2979	−3.2979

Hull and White give us $\mu = (5.09275, 6.50257, 7.33932, 8.05381)$, which gives us a short rate tree of

$n = 0$	$n = 1$	$n = 2$	$n = 3$
		10.6372	11.3517
	8.1515	8.9883	9.7028
5.0928	6.5026	7.3393	8.0538
	4.8536	5.6904	6.4049
		4.0414	4.7559

State prices and spot rates. We apply Duffie's formula to get

	$n = 0$	$n = 1$	$n = 2$	$n = 3$	n=4
			0.01805	0.03662	
		0.15839	0.19495	0.18772	
	1.00000	0.63357	0.46074	0.36842	
		0.15839	0.19817	0.19466	
			0.10865	0.03925	
	dis factor	0.95035	0.89056	0.82767	0.76388
	spot rate	5.0927	5.7954	6.3046	6.7335

Calibration. This is the answer. In general, we would observe spot rates in the data and use a numerical procedure to adjust μ until the spot rates in the model matched those in the data.

6 Final Thoughts

Hull and White provide a useful discrete-state approximation to the Vasicek model and a framework that can be applied to other models as well. Their method is relatively simple to apply and apparently works well in practice, but it's not the only way to approximate a continuous model. Nelson and Ramaswamy describe how binomial models might be designed to do similar things. Tauchen and Hussey (1991) go the other direction, and consider Markov chain approximations to continuous processes. And the math literature is filled with numerical approaches to solving "PDEs" (see Judd 1998). The key issues are the speed and simplicity of the calculations. What we'd like to see is a systematic assessment of various approaches with these criteria.

References

John Hull and Alan White, 1994 "Numerical procedures for implementing term structure models I: Single-factor models," *Journal of Derivatives*, Fall, 7-16.

John Hull and Alan White, 1996, "Using Hull-White interest rate trees," *Journal of Derivatives*, Spring, 26-36.

Kenneth Judd, 1998, *Numerical Methods in Economics*, MIT Press.

Daniel Nelson and Krishna Ramaswamy, 1990, "Simple binomial processes as diffusion approximations in financial models," *Review of Financial Studies* 3, 393-430.