The Black-Derman-Toy Model[a]

- This model is extensively used by practitioners.
- The BDT short rate process is the lognormal binomial interest rate process described on pp. 1002ff.[b]
- The volatility structure[c] is given by the market.
- From it, the short rate volatilities (thus v_i) are determined together with the baseline rates r_i.

---

[a]Black, Derman, & Toy (BDT) (1990), but essentially finished in 1986 according to Mehrling (2005).

[b]Repeated on next page.

[c]Recall Eq. (136) on p. 1053.

$$
\begin{tikzcd}
 & & & r_4 \\
 & & r_3 \arrow[ur] \arrow[dr] & \\
 & r_2 \arrow[ur] \arrow[dr] & & r_4 v_4 \\
r_1 \arrow[ur] \arrow[dr] & & r_3 v_3 \arrow[ur] \arrow[dr] & \\
 & r_2 v_2 \arrow[ur] \arrow[dr] & & r_4 v_4^2 \\
 & & r_3 v_3^2 \arrow[ur] \arrow[dr] & \\
 & & & r_4 v_4^3
\end{tikzcd}
$$

The Black-Derman-Toy Model (concluded)

- Our earlier binomial interest rate tree, in contrast, assumes v_i are given a priori.
- Lognormal models preclude negative short rates.

The BDT Model: Volatility Structure

- The volatility structure defines the yield volatilities of zero-coupon bonds of various maturities.
- Let the yield volatility of the i-period zero-coupon bond be denoted by κ_i.
- P_{u} is the price of the i-period zero-coupon bond one period from now if the short rate makes an up move.
- P_{d} is the price of the i-period zero-coupon bond one period from now if the short rate makes a down move.

The BDT Model: Volatility Structure (concluded)

- Corresponding to these two prices are the following yields to maturity,

$$y_{\mathrm{u}} \stackrel{\Delta}{=} P_{\mathrm{u}}^{-1/(i-1)} - 1,$$
$$y_{\mathrm{d}} \stackrel{\Delta}{=} P_{\mathrm{d}}^{-1/(i-1)} - 1.$$

- The yield volatility is defined as[a]

$$\kappa_i \stackrel{\Delta}{=} \frac{\ln(y_{\mathrm{u}}/y_{\mathrm{d}})}{2}.$$

---

[a]Ryecall Eq. (136) on p. 1053.

The BDT Model: Calibration

- The inputs to the BDT model are riskless zero-coupon bond yields and their volatilities.
- For economy of expression, all numbers are period based.
- Suppose inductively that we have calculated

$$(r_1, v_1), (r_2, v_2), \ldots, (r_{i-1}, v_{i-1}).$$

  - They define the binomial tree up to time $i-2$ (thus period $i-1$).[a]
  - Thus the spot rates up to time $i-1$ have been matched.

[a] Recall that (r_{i-1}, v_{i-1}) defines $i-1$ short rates at time $i-2$, which are applicable to period $i-1$.

The BDT Model: Calibration (continued)

- We now proceed to calculate r_i and v_i to extend the tree to cover period i.
- Assume the price of the i-period zero can move to P_{u} or P_{d} one period from now.
- Let y denote the current i-period spot rate, which is known.
- In a risk-neutral economy,
$$\frac{P_{\mathrm{u}} + P_{\mathrm{d}}}{2(1 + r_i)} = \frac{1}{(1 + y)^i}. \tag{155}$$
- Obviously, P_{u} and P_{d} are functions of the unknown r_i and v_i.

The BDT Model: Calibration (continued)

- Viewed from now, the future $(i-1)$-period spot rate at time 1 is uncertain.
- Recall that y_{u} and y_{d} represent the spot rates at the up node and the down node, respectively.[a]
- With κ_i^2 denoting their variance, we have

$$\kappa_i = \frac{1}{2} \ln \left(\frac{{P_{\mathrm{u}}}^{-1/(i-1)} - 1}{{P_{\mathrm{d}}}^{-1/(i-1)} - 1} \right). \tag{156}$$

[a] Recall p. 1162.

The BDT Model: Calibration (continued)

- Solving Eqs. (155)–(156) for r_i and v_i with backward induction takes $O(i^2)$ time.
  - That leads to a cubic-time algorithm.
- We next employ forward induction to derive a quadratic-time calibration algorithm.[a]
- Forward induction figures out, by moving *forward* in time, how much \$1 at a node contributes to the price.[b]
- This number is called the state price and is the price of the claim that pays \$1 at that node and zero elsewhere.

---

[a] W. J. Chen (R84526007) & Lyuu (1997); Lyuu (1999).
[b] Review p. 1030(a).

The BDT Model: Calibration (continued)

- Let the unknown baseline rate for period i be $r_i = r$.
- Let the unknown multiplicative ratio be $v_i = v$.
- Let the state prices at time $i-1$ be

$$P_1, P_2, \dots, P_i.$$

- They correspond to rates

$$r, rv, \dots, rv^{i-1}$$

  for period i, respectively.
- One dollar at time i has a present value of

$$f(r,v) \stackrel{\Delta}{=} \frac{P_1}{1+r} + \frac{P_2}{1+rv} + \frac{P_3}{1+rv^2} + \cdots + \frac{P_i}{1+rv^{i-1}}.$$

The BDT Model: Calibration (continued)

- By Eq. (156) on p. 1165, the yield volatility is

$$g(r,v) \triangleq \frac{1}{2} \ln \left(\frac{\left(\frac{P_{\mathrm{u},1}}{1+rv} + \frac{P_{\mathrm{u},2}}{1+rv^2} + \cdots + \frac{P_{\mathrm{u},i-1}}{1+rv^{i-1}} \right)^{-1/(i-1)} - 1}{\left(\frac{P_{\mathrm{d},1}}{1+r} + \frac{P_{\mathrm{d},2}}{1+rv} + \cdots + \frac{P_{\mathrm{d},i-1}}{1+rv^{i-2}} \right)^{-1/(i-1)} - 1} \right).$$

- Above, $P_{\mathrm{u},1}, P_{\mathrm{u},2}, \ldots$ denote the state prices at time $i-1$ *of* the subtree rooted at the up node.[a]
- And $P_{\mathrm{d},1}, P_{\mathrm{d},2}, \ldots$ denote the state prices at time $i-1$ *of* the subtree rooted at the down node.[b]

---

[a]Like $r_2 v_2$ on p. 1159.
[b]Like r_2 on p. 1159.

The BDT Model: Calibration (concluded)

- Note that every node maintains *three* state prices: $P_i, P_{\mathrm{u},i}, P_{\mathrm{d},i}$.
- Now solve

$$\begin{aligned} f(r,v) &= \frac{1}{(1+y)^i}, \\ g(r,v) &= \kappa_i, \end{aligned}$$

for $r = r_i$ and $v = v_i$.

- This $O(n^2)$-time algorithm appears on p. 382 of the textbook.

Calibrating the BDT Model with the Differential Tree (in seconds)[a]

Number of years	Running time	Number of years	Running time	Number of years	Running time
3000	398.880	39000	8562.640	75000	26182.080
6000	1697.680	42000	9579.780	78000	28138.140
9000	2539.040	45000	10785.850	81000	30230.260
12000	2803.890	48000	11905.290	84000	32317.050
15000	3149.330	51000	13199.470	87000	34487.320
18000	3549.100	54000	14411.790	90000	36795.430
21000	3990.050	57000	15932.370	120000	63767.690
24000	4470.320	60000	17360.670	150000	98339.710
27000	5211.830	63000	19037.910	180000	140484.180
30000	5944.330	66000	20751.100	210000	190557.420
33000	6639.480	69000	22435.050	240000	249138.210
36000	7611.630	72000	24292.740	270000	313480.390

75MHz Sun SPARCstation 20, one period per year.

[a]Lyuu (1999).

The BDT Model: Continuous-Time Limit

- The continuous-time limit of the BDT model is[a]

$$d \ln r = \left(\theta(t) + \frac{\sigma'(t)}{\sigma(t)} \ln r \right) dt + \sigma(t)\, dW.$$

- The short rate volatility $\sigma(t)$ should be a *declining* function of time for the model to display mean reversion.
  - That makes $\sigma'(t) < 0$.
- In particular, constant $\sigma(t)$ will not attain mean reversion.

[a]Jamshidian (1991).

The Black-Karasinski Model[a]

- The BK model stipulates that the short rate follows
$$d \ln r = \kappa(t)(\theta(t) - \ln r)\, dt + \sigma(t)\, dW.$$
- This explicitly mean-reverting model depends on time through $\kappa(\cdot)$, $\theta(\cdot)$, and $\sigma(\cdot)$.
- The BK model hence has one more degree of freedom than the BDT model.
- The speed of mean reversion $\kappa(t)$ and the short rate volatility $\sigma(t)$ are independent.

[a]Black & Karasinski (1991).

The Black-Karasinski Model: Discrete Time

- The discrete-time version of the BK model has the same representation as the BDT model.
- To maintain a combining binomial tree, however, requires some manipulations.
- The next plot illustrates the ideas in which

$$\begin{aligned} t_2 &\stackrel{\Delta}{=} t_1 + \Delta t_1, \\ t_3 &\stackrel{\Delta}{=} t_2 + \Delta t_2. \end{aligned}$$

$$
\begin{array}{ccccc}
 & & & \nearrow & \\
 & & \ln r_{\mathrm{d}}(t_2) & & \\
 & \nearrow & & \searrow & \\
\ln r(t_1) & & & & \ln r_{\mathrm{du}}(t_3) = \ln r_{\mathrm{ud}}(t_3) \\
 & \searrow & & \nearrow & \\
 & & \ln r_{\mathrm{u}}(t_2) & & \\
 & & & \searrow &
\end{array}
$$

The Black-Karasinski Model: Discrete Time (continued)

- Note that

$$\begin{aligned}\ln r_{\mathrm{d}}(t_2) &= \ln r(t_1) + \kappa(t_1)(\theta(t_1) - \ln r(t_1))\,\Delta t_1 - \sigma(t_1)\sqrt{\Delta t_1}\,,\\ \ln r_{\mathrm{u}}(t_2) &= \ln r(t_1) + \kappa(t_1)(\theta(t_1) - \ln r(t_1))\,\Delta t_1 + \sigma(t_1)\sqrt{\Delta t_1}\,.\end{aligned}$$

- To make sure an up move followed by a down move coincides with a down move followed by an up move,

$$\begin{aligned}&\ln r_{\mathrm{d}}(t_2) + \kappa(t_2)(\theta(t_2) - \ln r_{\mathrm{d}}(t_2))\,\Delta t_2 + \sigma(t_2)\sqrt{\Delta t_2}\,,\\ =\;&\ln r_{\mathrm{u}}(t_2) + \kappa(t_2)(\theta(t_2) - \ln r_{\mathrm{u}}(t_2))\,\Delta t_2 - \sigma(t_2)\sqrt{\Delta t_2}\,.\end{aligned}$$

The Black-Karasinski Model: Discrete Time (continued)

- They imply

$$\kappa(t_2) = \frac{1 - (\sigma(t_2)/\sigma(t_1))\sqrt{\Delta t_2/\Delta t_1}}{\Delta t_2}. \tag{157}$$

- So from Δt_1, we can calculate the Δt_2 that satisfies the combining condition and then iterate.
  - $t_0 \to \Delta t_1 \to t_1 \to \Delta t_2 \to t_2 \to \Delta t_3 \to \cdots \to T$ (roughly).[a]

[a] As $\kappa(t), \theta(t), \sigma(t)$ are independent of r, the Δt_i will not depend on r either.

The Black-Karasinski Model: Discrete Time (concluded)

- Unequal durations Δt_i are often necessary to ensure a combining tree.[a]

---

[a] Amin (1991); C. I. Chen (`R98922127`) (2011); Lok (`D99922028`) & Lyuu (2016, 2017).

Problems with Lognormal Models in General

- Lognormal models such as BDT and BK share the problem that $E^{\pi}[\,M(t)\,] = \infty$ for any finite t if they model the continuously compounded rate.[a]
- So periodically compounded rates should be modeled.[b]
- Another issue is computational.
- Lognormal models usually do not admit of analytical solutions to even basic fixed-income securities.
- As a result, to price short-dated derivatives on long-term bonds, the tree has to be built over the life of the underlying asset instead of the life of the derivative.

[a] Hogan & Weintraub (1993).
[b] Sandmann & Sondermann (1993).

Problems with Lognormal Models in General (concluded)

- This problem can be somewhat mitigated by adopting variable-duration time steps.[a]
  - Use a fine time step up to the maturity of the short-dated derivative.
  - Use a coarse time step beyond the maturity.
- A down side of this procedure is that it has to be tailor-made for each derivative.
- Finally, empirically, interest rates do not follow the lognormal distribution.

[a]Hull & White (1993).

The Extended Vasicek Model[a]

- Hull and White proposed models that extend the Vasicek model and the CIR model.
- They are called the extended Vasicek model and the extended CIR model.
- The extended Vasicek model adds time dependence to the original Vasicek model,

$$dr = (\theta(t) - a(t)\, r)\, dt + \sigma(t)\, dW.$$

- Like the Ho-Lee model, this is a normal model.
- The inclusion of $\theta(t)$ allows for an exact fit to the current spot rate curve.

[a] Hull & White (1990).

The Extended Vasicek Model (concluded)

- Function $\sigma(t)$ defines the short rate volatility, and $a(t)$ determines the shape of the volatility structure.
- Many European-style securities can be evaluated analytically.
- Efficient numerical procedures can be developed for American-style securities.

The Hull-White Model

- The Hull-White model is the following special case,

$$dr = (\theta(t) - ar)\, dt + \sigma\, dW. \tag{158}$$

- When the current term structure is matched,[a]

$$\theta(t) = \frac{\partial f(0,t)}{\partial t} + af(0,t) + \frac{\sigma^2}{2a}\left(1 - e^{-2at}\right).$$

  - Recall that $f(0,t)$ defines the forward rate curve.

---

[a]Hull & White (1993).

The Extended CIR Model

- In the extended CIR model the short rate follows

$$dr = (\theta(t) - a(t)\, r)\, dt + \sigma(t)\sqrt{r}\, dW.$$

- The functions $\theta(t)$, $a(t)$, and $\sigma(t)$ are implied from market observables.
- With constant parameters, there exist analytical solutions to a small set of interest rate-sensitive securities.

The Hull-White Model: Calibration[a]

- We describe a trinomial forward induction scheme to calibrate the Hull-White model given a and σ.
- As with the Ho-Lee model, the set of achievable short rates is evenly spaced.
- Let r_0 be the annualized, continuously compounded short rate at time zero.
- Every short rate on the tree takes on a value

$$r_0 + j\Delta r$$

for some integer j.

[a]Hull & White (1993).

The Hull-White Model: Calibration (continued)

- Time increments on the tree are also equally spaced at Δt apart.
- Hence nodes are located at times $i\Delta t$ for $i = 0, 1, 2, \ldots$.
- We shall refer to the node on the tree with

$$t_i \stackrel{\Delta}{=} i\Delta t,$$
$$r_j \stackrel{\Delta}{=} r_0 + j\Delta r,$$

  as the (i, j) node.
- The short rate at node (i, j), which equals r_j, is effective for the time period $[\, t_i, t_{i+1})$.

The Hull-White Model: Calibration (continued)

- Use

$$\mu_{i,j} \stackrel{\Delta}{=} \theta(t_i) - a r_j \tag{159}$$

  to denote the drift rate[a] of the short rate as seen from node (i, j).

- The three distinct possibilities for node (i, j) with three branches incident from it are displayed on p. 1187.

- The middle branch may be an increase of Δr, no change, or a decrease of Δr.

[a]Or, the annualized expected change.

The Hull-White Model: Calibration (continued)

$$\begin{tikzcd}
 & (i+1, j+2) \\
 & (i+1, j+1) \\
(i, j) \arrow[ruu] \arrow[ru] \arrow[r] & (i+1, j)
\end{tikzcd}$$

$$\begin{tikzcd}
 & (i+1, j+1) \\
(i, j) \arrow[ru] \arrow[r] \arrow[rd] & (i+1, j) \\
 & (i+1, j-1)
\end{tikzcd}$$

$$\begin{tikzcd}
(i, j) \arrow[r] \arrow[rd] \arrow[rdd] & (i+1, j) \\
 & (i+1, j-1) \\
 & (i+1, j-2)
\end{tikzcd}$$

The Hull-White Model: Calibration (continued)

- The upper and the lower branches bracket the middle branch.
- Define

$$
\begin{aligned}
p_1(i,j) &\stackrel{\Delta}{=} \text{the probability of following the upper branch from node } (i,j),\\
p_2(i,j) &\stackrel{\Delta}{=} \text{the probability of following the middle branch from node } (i,j),\\
p_3(i,j) &\stackrel{\Delta}{=} \text{the probability of following the lower branch from node } (i,j).
\end{aligned}
$$

- The root of the tree is set to the current short rate r_0.
- Inductively, the drift $\mu_{i,j}$ at node (i,j) is a function of (the still unknown) $\theta(t_i)$.
  - It describes the expected change from node (i,j).

The Hull-White Model: Calibration (continued)

- Once $\theta(t_i)$ is available, $\mu_{i,j}$ can be derived via Eq. (159) on p. 1186.
- This in turn determines the branching scheme at every node (i, j) for each j, as we will see shortly.
- The value of $\theta(t_i)$ must thus be made consistent with the spot rate $r(0, t_{i+2})$.[a]

---

[a] Not $r(0, t_{i+1})$!

The Hull-White Model: Calibration (continued)

- The branches emanating from node (i, j) with their probabilities[a] must be chosen to be consistent with $\mu_{i,j}$ and σ.
- This is done by selecting the middle node to be as closest to the current short rate r_j plus the drift $\mu_{i,j} \Delta t$.[b]

---

[a] That is, $p_1(i, j)$, $p_2(i, j)$, and $p_3(i, j)$.

[b] A precursor of Lyuu and C. Wu's (R90723065) (2003, 2005) mean-tracking idea, which is the precursor of the binomial-trinomial tree of Dai (B82506025, R86526008, D8852600) & Lyuu (2006, 2008, 2010).

The Hull-White Model: Calibration (continued)

- Let k be the number among $\{ j-1, j, j+1 \}$ that makes the short rate reached by the middle branch, r_k, closest to

$$r_j + \mu_{i,j} \Delta t.$$

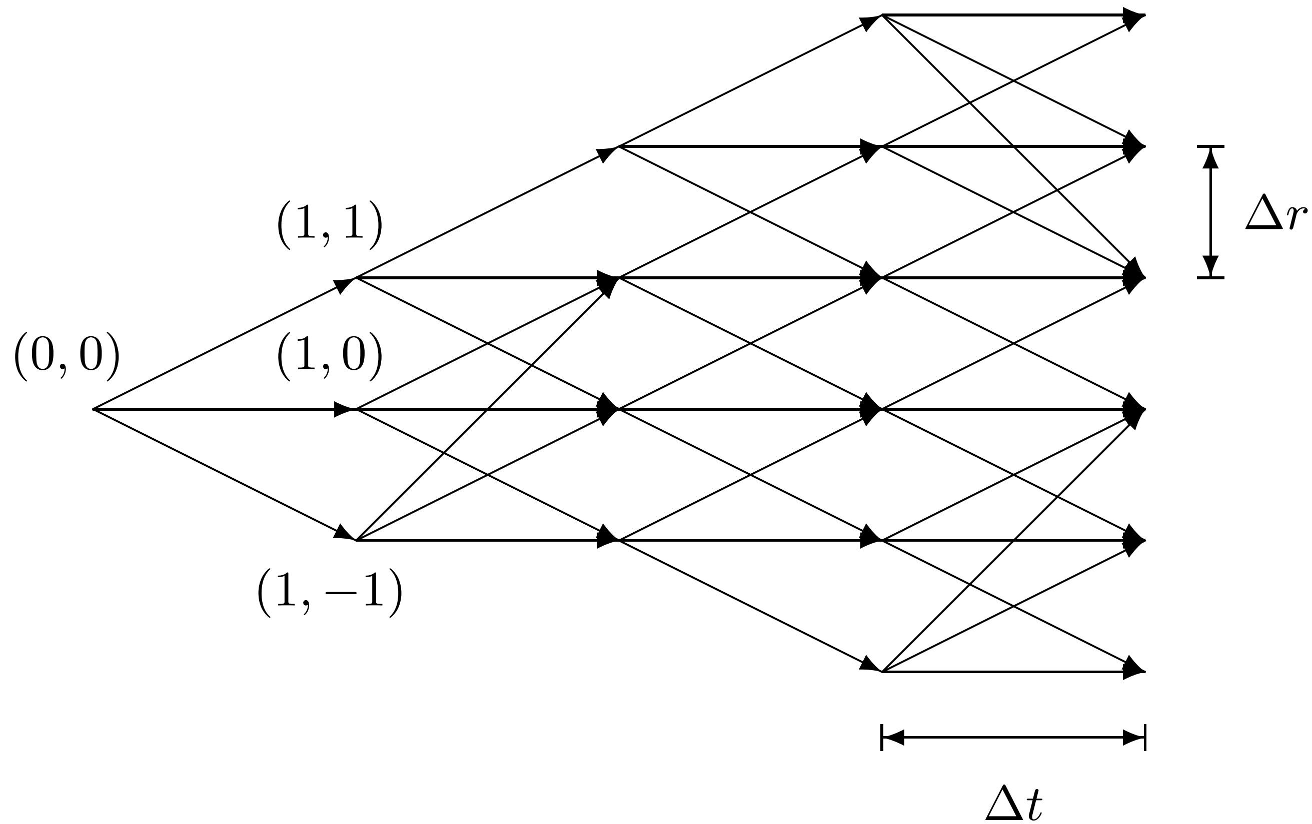
  – But note that $\mu_{i,j}$ is still *not* computed yet.

- Then the three nodes following node (i, j) are nodes

$$(i+1, k+1), (i+1, k), (i+1, k-1).$$

- See p. 1192 for a possible geometry.

- The resulting tree combines.



$(0,0)$
$(1,1)$
$(1,0)$
$(1,-1)$
Δr
Δt

The Hull-White Model: Calibration (continued)

- The probabilities for moving along these branches are functions of $\mu_{i,j}$, σ, j, and k:

$$p_1(i,j) = \frac{\sigma^2 \Delta t + \eta^2}{2(\Delta r)^2} + \frac{\eta}{2\Delta r}, \tag{160}$$

$$p_2(i,j) = 1 - \frac{\sigma^2 \Delta t + \eta^2}{(\Delta r)^2}, \tag{160$'$}$$

$$p_3(i,j) = \frac{\sigma^2 \Delta t + \eta^2}{2(\Delta r)^2} - \frac{\eta}{2\Delta r}, \tag{160$''$}$$

where

$$\eta \stackrel{\Delta}{=} \mu_{i,j} \Delta t + (j-k)\, \Delta r.$$

The Hull-White Model: Calibration (continued)

- As trinomial tree algorithms are but explicit methods in disguise,[a] certain relations must hold for Δr and Δt to guarantee stability.
- It can be shown that their values must satisfy

$$\frac{\sigma\sqrt{3\Delta t}}{2} \leq \Delta r \leq 2\sigma\sqrt{\Delta t}$$

for the probabilities to lie between zero and one.

  - For example, Δr can be set to $\sigma\sqrt{3\Delta t}$.[b]

- Now it only remains to determine $\theta(t_i)$.

---

[a] Recall p. 826.

[b] Hull & White (1988).

The Hull-White Model: Calibration (continued)

- At this point at time t_i,

$$r(0, t_1), r(0, t_2), \dots, r(0, t_{i+1})$$

  have already been matched.

- Let $Q(i, j)$ be the state price at node (i, j).
- By construction, the state prices $Q(i, j)$ for all j are known by now.
- We begin with state price $Q(0, 0) = 1$.

The Hull-White Model: Calibration (continued)

- Let $\hat{r}(i)$ refer to the short rate value at time t_i.
- The value at time zero of a zero-coupon bond maturing at time t_{i+2} is then

$$
\begin{aligned}
& e^{-r(0,t_{i+2})(i+2)\,\Delta t} \\
= \; & \sum_j Q(i,j)\, e^{-r_j \Delta t}\, E^{\pi}\left[e^{-\hat{r}(i+1)\,\Delta t} \,\middle|\, \hat{r}(i) = r_j \right]. \qquad (161)
\end{aligned}
$$

- The right-hand side represents the value of \$1 at time t_{i+2} as seen at node (i,j) at time[a] t_i before being discounted by $Q(i,j)$.

---

[a] Thus $\hat{r}(i+1)$ is stochastic.

The Hull-White Model: Calibration (continued)

- The expectation in Eq. (161) can be approximated by[a]

$$\begin{aligned} & E^{\pi}\left[\, e^{-\hat{r}(i+1)\,\Delta t} \,\middle|\, \hat{r}(i) = r_j \right] \\ \approx\ & e^{-r_j \Delta t}\left(1 - \mu_{i,j}(\Delta t)^2 + \frac{\sigma^2 (\Delta t)^3}{2}\right). \qquad (162) \end{aligned}$$

  – This solves the chicken-egg problem!

- Substitute Eq. (162) into Eq. (161) and replace $\mu_{i,j}$ with $\theta(t_i) - a r_j$ to obtain

$$\theta(t_i) \approx \frac{\sum_j Q(i,j)\, e^{-2 r_j \Delta t}\left(1 + a r_j (\Delta t)^2 + \sigma^2 (\Delta t)^3/2\right) - e^{-r(0,t_{i+2})(i+2)\,\Delta t}}{(\Delta t)^2 \sum_j Q(i,j)\, e^{-2 r_j \Delta t}}.$$

[a] See Exercise 26.4.2 of the textbook.

The Hull-White Model: Calibration (continued)

- For the Hull-White model, the expectation in Eq. (162) is actually known analytically by Eq. (30) on p. 179:

$$
\begin{aligned}
& E^{\pi}\left[e^{-\hat{r}(i+1)\,\Delta t} \,\middle|\, \hat{r}(i) = r_j \right] \\
= \quad & e^{-r_j \Delta t + (-\theta(t_i) + a r_j + \sigma^2 \Delta t/2)(\Delta t)^2}.
\end{aligned}
$$

- Therefore, alternatively,

$$
\theta(t_i) = \frac{r(0, t_{i+2})(i+2)}{\Delta t} + \frac{\sigma^2 \Delta t}{2} + \frac{\ln \sum_j Q(i,j)\, e^{-2 r_j \Delta t + a r_j (\Delta t)^2}}{(\Delta t)^2}.
$$

- With $\theta(t_i)$ in hand, we can compute $\mu_{i,j}$.[a]

[a] See Eq. (159) on p. 1186.

The Hull-White Model: Calibration (concluded)

- With $\mu_{i,j}$ available, we compute the probabilities.[a]
- Finally the state prices at time t_{i+1}:

$$Q(i+1,j) = \sum_{(i,j^*) \text{ is connected to } (i+1,j) \text{ with probability } p_{j^*}} p_{j^*} e^{-r_{j^*}\Delta t} Q(i,j^*).$$

- There are at most 5 choices for j^* (why?).
- The total running time is $O(n^2)$.
- The space requirement is $O(n)$ (why?).

[a] See Eqs. (160) on p. 1193.

Comments on the Hull-White Model

- One can try different values of a and σ for each option.
- Or have an a value common to all options but use a different σ value for each option.
- Either approach can match all the option prices exactly.
- But suppose the demand is for a single set of parameters that replicate *all* option prices.
- Then the Hull-White model can be calibrated to all the observed option prices by choosing a and σ that minimize the mean-squared pricing error.[a]

[a] Hull & White (1995).

The Hull-White Model: Calibration with Irregular Trinomial Trees

- The previous calibration algorithm is quite general.
- For example, it can be modified to apply to cases where the diffusion term has the form σr^b.
- But it has at least two shortcomings.
- First, the resulting trinomial tree is irregular (p. 1192).
  - So it is harder to program (for nonprogrammers).
- The second shortcoming is a consequence of the tree's irregular shape.

The Hull-White Model: Calibration with Irregular Trinomial Trees (concluded)

- Recall that the algorithm figured out $\theta(t_i)$ that matches the spot rate $r(0, t_{i+2})$ in order to determine the branching schemes for the nodes at time t_i.
- But without those branches, the tree was not specified, and backward induction on the tree was not possible.
- To avoid this chicken-egg dilemma, the algorithm turned to the continuous-time model to evaluate Eq. (161) on p. 1196 that helps derive $\theta(t_i)$.
- The resulting $\theta(t_i)$ hence might not yield a tree that matches the spot rates exactly.

The Hull-White Model: Calibration with Regular Trinomial Trees[a]

- The next, simpler algorithm exploits the fact that the Hull-White model has a constant diffusion term σ.
- The resulting trinomial tree will be regular.
- All the $\theta(t_i)$ terms can be chosen by backward induction to match the spot rates exactly.
- The tree is constructed in two phases.

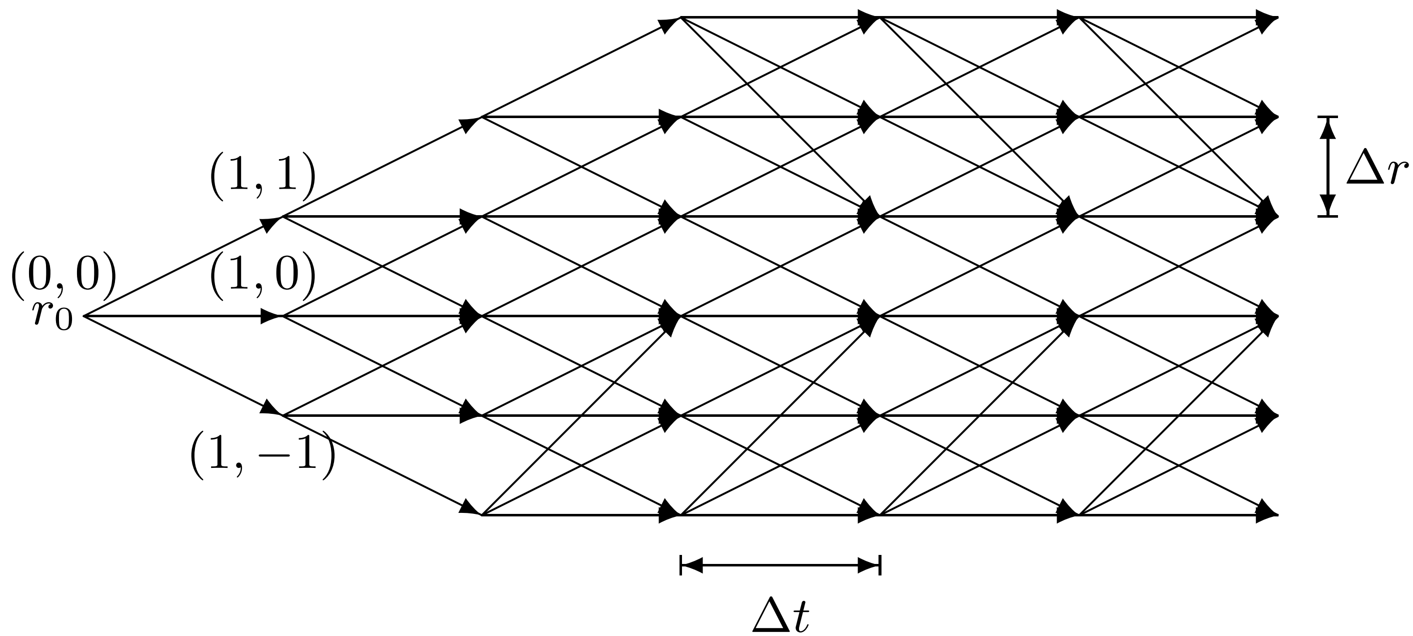
[a] Hull & White (1994).

The Hull-White Model: Calibration with Regular Trinomial Trees (continued)

- In the first phase, a tree is built for the $\theta(t) = 0$ case, which is an Ornstein-Uhlenbeck process:

$$dr = -ar\,dt + \sigma\,dW, \quad r(0) = 0.$$

  - The tree is dagger-shaped (see p. 1205).
  - The number of nodes above the r_0-line is $j_{\max}$, and that below the line is $j_{\min}$.
  - They will be picked so that the probabilities (160) on p. 1193 are positive for all nodes.



The short rate at node $(0,0)$ equals $r_0 = 0$; here $j_{\max} = 3$ and $j_{\min} = 2$.

The Hull-White Model: Calibration with Regular Trinomial Trees (concluded)

- The tree's branches and probabilities are now in place.
- Phase two fits the term structure.
  - Backward induction is applied to calculate the β_i to *add* to the short rates on the tree at time t_i so that the spot rate $r(0, t_{i+1})$ is matched.[a]

---

[a]Contrast this with the previous algorithm, where it was $r(0, t_{i+2})$ that was matched!

The Hull-White Model: Calibration

- Set $\Delta r = \sigma\sqrt{3\Delta t}$ and assume that $a > 0$.
- Node (i, j) is a top node if $j = j_{\max}$ and a bottom node if $j = -j_{\min}$.
- Because the root of the tree has a short rate of $r_0 = 0$, phase one adopts $r_j = j\Delta r$.
- Hence the probabilities in Eqs. (160) on p. 1193 use

$$\eta \stackrel{\Delta}{=} -aj\Delta r\Delta t + (j-k)\,\Delta r.$$

- Recall that k tracks the middle branch.

The Hull-White Model: Calibration (continued)

- The probabilities become

$$
\begin{aligned}
& p_1(i, j) \\
= \quad & \frac{1}{6} + \frac{a^2 j^2 (\Delta t)^2 - 2aj\Delta t(j-k) + (j-k)^2 - aj\Delta t + (j-k)}{2}, \qquad (163) \\
& p_2(i, j) \\
= \quad & \frac{2}{3} - \left[\, a^2 j^2 (\Delta t)^2 - 2aj\Delta t(j-k) + (j-k)^2 \,\right], \qquad (164) \\
& p_3(i, j) \\
= \quad & \frac{1}{6} + \frac{a^2 j^2 (\Delta t)^2 - 2aj\Delta t(j-k) + (j-k)^2 + aj\Delta t - (j-k)}{2}. \qquad (165)
\end{aligned}
$$

- p_1: up move; p_2: flat move; p_3: down move.

The Hull-White Model: Calibration (continued)

- The dagger shape dictates this:
  - Let $k = j - 1$ if node (i, j) is a top node.
  - Let $k = j + 1$ if node (i, j) is a bottom node.
  - Let $k = j$ for the rest of the nodes.
- Note that the probabilities are identical for nodes (i, j) with the same j.
- Furthermore, $p_1(i, j) = p_3(i, -j)$.

The Hull-White Model: Calibration (continued)

- The inequalities

$$\frac{3-\sqrt{6}}{3} < ja\Delta t < \sqrt{\frac{2}{3}} \tag{166}$$

ensure that all the branching probabilities are positive in the upper half of the tree, that is, $j > 0$ (verify this).

- Similarly, the inequalities

$$-\sqrt{\frac{2}{3}} < ja\Delta t < -\frac{3-\sqrt{6}}{3}$$

ensure that the probabilities are positive in the lower half of the tree, that is, $j < 0$.

The Hull-White Model: Calibration (continued)

- To further make the tree symmetric across the r_0-line, we let $j_{\min} = j_{\max}$.
- As
$$\frac{3-\sqrt{6}}{3} \approx 0.184,$$
a good choice is
$$j_{\max} = \lceil 0.184/(a\Delta t) \rceil.$$

The Hull-White Model: Calibration (continued)

- Phase two computes the β_is to fit the spot rates.
- We begin with state price $Q(0,0) = 1$.
- Inductively, suppose that spot rates

$$r(0, t_1), r(0, t_2), \ldots, r(0, t_i)$$

have already been matched.

- By construction, the state prices $Q(i, j)$ for all j are known by now.

The Hull-White Model: Calibration (continued)

- The value of a zero-coupon bond maturing at time t_{i+1} equals

$$e^{-r(0,t_{i+1})(i+1)\,\Delta t} = \sum_j Q(i,j)\, e^{-(\beta_i + r_j)\Delta t}$$

  by risk-neutral valuation.

- Hence

$$\beta_i = \frac{r(0,t_{i+1})(i+1)\,\Delta t + \ln \sum_j Q(i,j)\, e^{-r_j\,\Delta t}}{\Delta t}. \tag{167}$$

The Hull-White Model: Calibration (concluded)

- The short rate at node (i, j) now equals $\beta_i + r_j$.
- The state prices at time t_{i+1},

$$Q(i+1, j), \quad -\min(i+1, j_{\max}) \leq j \leq \min(i+1, j_{\max}),$$

can now be calculated as before.[a]

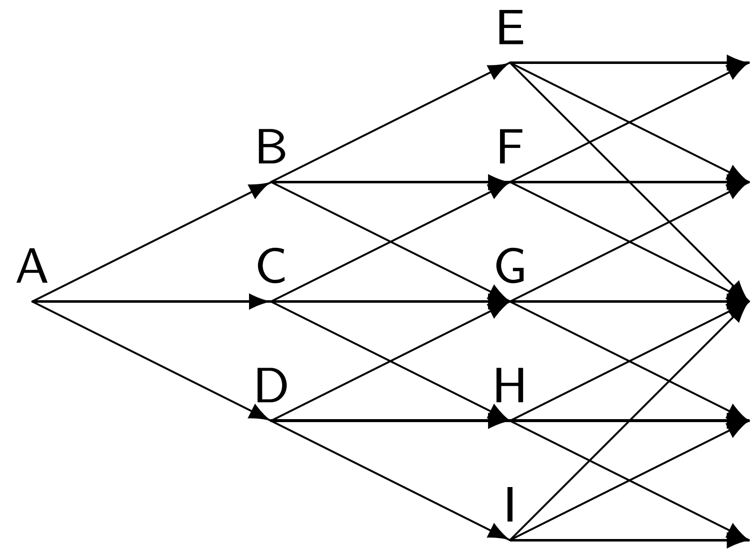
- The total running time is $O(nj_{\max})$.
- The space requirement is $O(n)$.

---

[a]Recall p. 1199.

A Numerical Example

- Assume $a = 0.1$, $\sigma = 0.01$, and $\Delta t = 1$ (year).
- Immediately, $\Delta r = 0.0173205$ and $j_{\max} = 2$.
- The plot on p. 1216 illustrates the 3-period trinomial tree after phase one.
- For example, the branching probabilities for node E are calculated by Eqs. (163)–(165) on p. 1208 with $j = 2$ and $k = 1$.



Node	A, C, G	B, F	E	D, H	I
r (%)	0.00000	1.73205	3.46410	-1.73205	-3.46410
p_1	0.16667	0.12167	0.88667	0.22167	0.08667
p_2	0.66667	0.65667	0.02667	0.65667	0.02667
p_3	0.16667	0.22167	0.08667	0.12167	0.88667

A Numerical Example (continued)

- Suppose that phase two is to fit the spot rate curve

$$0.08 - 0.05 \times e^{-0.18 \times t}.$$

- The annualized continuously compounded spot rates are

$$r(0,1) = 3.82365\%, r(0,2) = 4.51162\%, r(0,3) = 5.08626\%.$$

- Start with state price $Q(0,0) = 1$ at node A.

A Numerical Example (continued)

- Now, by Eq. (167) on p. 1213,

$$\beta_0 = r(0,1) + \ln Q(0,0)\, e^{-r_0} = r(0,1) = 3.82365\%.$$

- Hence the short rate at node A equals

$$\beta_0 + r_0 = 3.82365\%.$$

- The state prices at year one are calculated as

$$\begin{aligned} Q(1,1) &= p_1(0,0)\, e^{-(\beta_0 + r_0)} = 0.160414, \\ Q(1,0) &= p_2(0,0)\, e^{-(\beta_0 + r_0)} = 0.641657, \\ Q(1,-1) &= p_3(0,0)\, e^{-(\beta_0 + r_0)} = 0.160414. \end{aligned}$$

A Numerical Example (continued)

- The 2-year rate spot rate $r(0,2)$ is matched by picking

$$\beta_1 = r(0,2)\times 2 + \ln\left[\, Q(1,1)\, e^{-\Delta r} + Q(1,0) + Q(1,-1)\, e^{\Delta r} \,\right] = 5.20459\%.$$

- Hence the short rates at nodes B, C, and D equal

$$\beta_1 + r_j,$$

  where $j = 1, 0, -1$, respectively.

- They are found to be 6.93664%, 5.20459%, and 3.47254%.

A Numerical Example (continued)

- The state prices at year two are calculated as

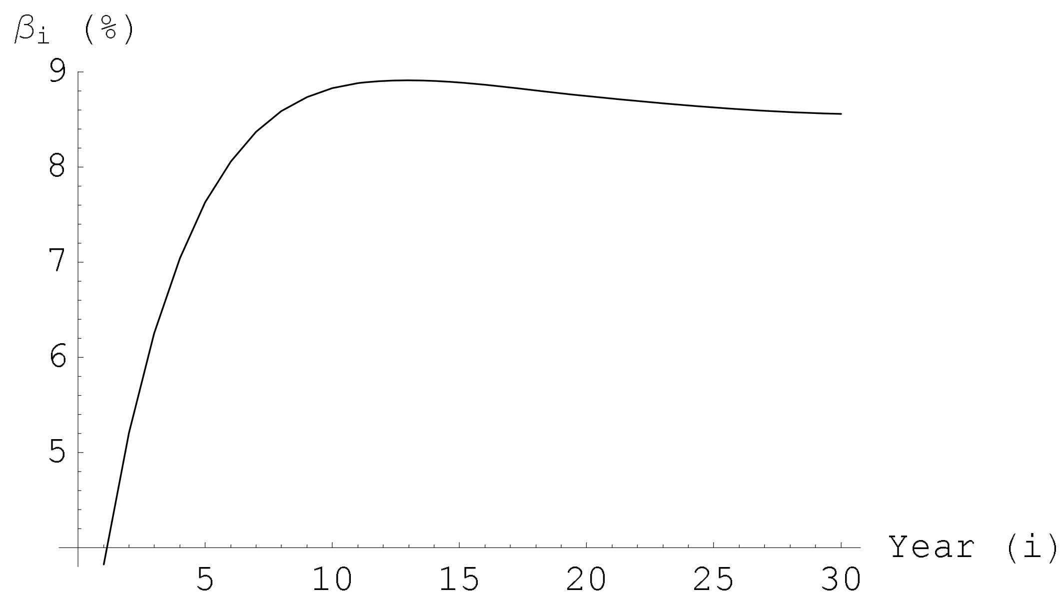
$$
\begin{aligned}
Q(2,2) &= p_1(1,1)\, e^{-(\beta_1+r_1)} Q(1,1) = 0.018209,\\
Q(2,1) &= p_2(1,1)\, e^{-(\beta_1+r_1)} Q(1,1) + p_1(1,0)\, e^{-(\beta_1+r_0)} Q(1,0)\\
&= 0.199799,\\
Q(2,0) &= p_3(1,1)\, e^{-(\beta_1+r_1)} Q(1,1) + p_2(1,0)\, e^{-(\beta_1+r_0)} Q(1,0)\\
&\quad + p_1(1,-1)\, e^{-(\beta_1+r_{-1})} Q(1,-1) = 0.473597,\\
Q(2,-1) &= p_3(1,0)\, e^{-(\beta_1+r_0)} Q(1,0) + p_2(1,-1)\, e^{-(\beta_1+r_{-1})} Q(1,-1)\\
&= 0.203263,\\
Q(2,-2) &= p_3(1,-1)\, e^{-(\beta_1+r_{-1})} Q(1,-1) = 0.018851.
\end{aligned}
$$

A Numerical Example (concluded)

- The 3-year rate spot rate $r(0,3)$ is matched by picking

$$\begin{aligned}\beta_2 \;=\;& r(0,3)\times 3+\ln\Big[\,Q(2,2)\,e^{-2\times\Delta r}+Q(2,1)\,e^{-\Delta r}+Q(2,0)\\ &+Q(2,-1)\,e^{\Delta r}+Q(2,-2)\,e^{2\times\Delta r}\Big]=6.25359\%.\end{aligned}$$

- Hence the short rates at nodes E, F, G, H, and I equal β_2+r_j, where $j=2,1,0,-1,-2$, respectively.
- They are found to be 9.71769%, 7.98564%, 6.25359%, 4.52154%, and 2.78949%.
- The figure on p. 1222 plots β_i for $i=0,1,\ldots,29$.



β_i (%)
9
8
7
6
5
5
10
15
20
25
30
Year (i)

The (Whole) Yield Curve Approach

- We have seen several Markovian short rate models.
- The Markovian approach is computationally efficient.
- But it is difficult to model the behavior of yields and bond prices of different maturities.
- The alternative yield curve approach regards the whole term structure as the state of a process and directly specifies how it evolves.

The Heath-Jarrow-Morton (HJM) Model[a]

- This influential model is a forward rate model.
- The HJM model specifies the initial forward rate curve and the forward rate volatility structure.
  - The volatility structure describes the volatility of each forward rate for a given maturity date.
- Like the Black-Scholes option pricing model, neither risk preference assumptions nor the drifts of forward rates are needed.

[a]Heath, Jarrow, & Morton (1992).

The HJM Model (continued)

- Within a finite-time horizon $[\,0, U\,]$, we take as given the time-zero forward rate curve $f(0, T)$ for $T \in [\,0, U\,]$.
- Since this curve is used as the boundary value at $t = 0$, perfect fit to the observed term structure is *automatic*.
- The forward rates are driven by k stochastic factors.

The HJM Model (continued)

- Specifically the forward rate movements are governed by the stochastic process,

$$df(t,T) = \mu(t,T)\,dt + \sum_{i=1}^{k} \sigma_i(t,T)\,dW_i, \tag{168}$$

where μ and σ_i may depend on the past history of the independent Wiener processes $W_1, W_2, \ldots, W_k$.

- One-factor models seem to perform better than multifactor models empirically, at least for *pricing* short-dated options.[a]

[a] Amin & Morton (1994).

The HJM Model (continued)

- But two-factor models perform better in *hedging* caps and floors.[a]
- Kamakura (2019) has a 10-factor[b] (14-factor[c]) HJM model for the U.S. Treasuries (German bonds, respectively).
- A unique equivalent martingale measure π can be established under which the prices of interest rate derivatives do not depend on the market prices of risk.

---

[a] Gupta & Subrahmanyam (2001, 2005).

[b] See `http://www.kamakuraco.com/KamakuraReleasesNewStochasticVolatilityModel`

[c] See `http://www.kamakuraco.com/KamakuraReleases14FactorHeathJarrowandMorto.`

The HJM Model (continued)

Theorem 22 (1) *For all* $0 < t \leq T$,

$$\mu(t,T) = \sum_{i=1}^{k} \sigma_i(t,T) \int_t^T \sigma_i(t,u)\, du \tag{169}$$

holds under π *almost surely.* (2) *The bond price dynamics under* π *is given by*

$$\frac{dP(t,T)}{P(t,T)} = r(t)\, dt - \sum_{i=1}^{k} \sigma_{\mathrm{p},i}(t,T)\, dW_i, \tag{170}$$

where $\sigma_{\mathrm{p},i}(t,T) \equiv \int_t^T \sigma_i(t,u)\, du$.

The HJM Model (concluded)

- Hence choosing the volatility functions $\sigma_i(t, T)$ of the forward rate dynamics under π uniquely determines the drift parameters under π and the prices of all claims.

The Use of the HJM Model

- Take the one-factor model,

$$df(t,T) = \mu(t,T)\,dt + \sigma(t,T)\,dW_t.$$

- To use the HJM model, we first pick $\sigma(t,T)$.
- This is the modeling part.
- The drift parameters are then determined by Eq. (169) on p. 1228.
- Now fetch today's forward rate curve $\{ f(0,T), T \geq 0 \}$ and integrate it to obtain the forward rates,

$$f(t,T) = f(0,T) + \int_0^t \mu(s,T)\,ds + \int_0^t \sigma(s,T)\,dW_s.$$

The Use of the HJM Model (concluded)

- Compute the future bond prices by

$$P(t,T) = e^{-\int_t^T f(t,s)\,ds}$$

  if necessary.

- European-style derivatives can be priced by simulating many paths and taking average.

Short Rate under the HJM Model

- From Eq. (26.19) of the textbook, the short rate follows the following SDE,

$$\begin{aligned} dr(t) &= \frac{\partial f(0,t)}{\partial t}\, dt \\ &\quad + \left[\int_0^t \left(\sigma_{\mathrm{p}}(s,t)\, \frac{\partial \sigma(s,t)}{\partial t} + \sigma(s,t)^2 \right) ds \right] dt \\ &\quad + \left(\int_0^t \frac{\partial \sigma(s,t)}{\partial t}\, dW_s \right) dt + \sigma(t,t)\, dW_t. \quad (171) \end{aligned}$$

- Since the second and the third terms on the right-hand side depend on the history of σ_{p} and/or dW, they can make r non-Markovian.

Short Rate under the HJM Model (concluded)

- If $\sigma_{\mathrm{p}}(t, T) = \sigma(T - t)$ for a constant σ, the short rate process r becomes Markovian.
- Then Eq. (171) on p. 1232 is reduced to

$$dr = \left(\frac{\partial f(0, t)}{\partial t} + \sigma^2 t \right) dt + \sigma \, dW.$$

- This is the continuous-time Ho-Lee model (154) on p. 1154.[a]
- See Carverhill (1994) and Jeffrey (1995) for conditions for the short rate to be Markovian.

[a]See p. 392 of the textbook.

The Alternative HJM Model

- Alternatively, we can start with the bond process under π:

$$\frac{dP(t,T)}{P(t,T)} = r(t)\,dt + \sum_{i=1}^{k} \sigma_{\mathrm{p},i}(t,T)\,dW_i. \tag{172}$$

- Then[a]

$$\begin{aligned} df(t,T) &= \sum_{i=1}^{k} \sigma_{\mathrm{p},i}(t,T)\,\frac{\partial \sigma_{\mathrm{p},i}(t,T)}{\partial T}\,dt \\ &\quad -\sum_{i=1}^{k} \frac{\partial \sigma_{\mathrm{p},i}(t,T)}{\partial T}\,dW_i. \end{aligned}$$

[a] Carverhill (1995); Musiela & Rutkowski (1997); Hull (1999).

Gaussian HJM Models[a]

- A nonstochastic volatility depends on only t and T.
- When the forward rate volatilities $\sigma_i(t, T)$ are nonstochastic, we have a Gaussian HJM model.
- For Gaussian HJM models, the bond price volatilities $\sigma_{\mathrm{p},i}(t, T)$ must also be nonstochastic.
- The forward rates have a normal distribution, whereas the bond prices have a lognormal distribution.

[a] Musiela & Rutkowski (1997).

Gaussian HJM Models (concluded)

- $\sigma(t, T) = \sigma$: The Ho-Lee model (154) on p. 1154 obtains.
- $\sigma(t, T) = \sigma e^{-a(T-t)}$: The Hull-White model (158) on p. 1182 obtains.
- $\sigma(t, T) = \sigma_0 + \sigma_1(T - t)$: The linear absolute model.[a]
- $\sigma(t, T) = \sigma\,[\,\gamma(T - t) + 1\,]\,e^{-(\lambda/2)(T-t)}$: The Mercurio-Moraleda (2000) model.

---

[a] Gupta & Subrahmanyam (2001, 2005).

Local-Volatility HJM Models[a]

- If the forward rate volatilities $\sigma_i(t, T, f(t, T))$ depend on t, T, and $f(t, T)$ only, we have a local-volatility HJM model.
- The same term may also apply to HJM models whose bond price volatilities $\sigma_{\mathrm{p},i}(t, T, P(t, T))$ depend on t, T, and $P(t, T)$ only.

[a]Brigo & Mercurio (2006).

Local-Volatility HJM Models (continued)

- The (nearly) proportional volatility model:[a]

$$\sigma(t,T,f(t,T)) = \sigma_0 \min(\kappa, f(t,T)), \quad \sigma_0, \kappa > 0.$$

- The proportional volatility model:[b]

$$\sigma(t,T,f(t,T)) = \sigma_0 f(t,T). \tag{173}$$

- The linear proportional model:[c]

$$\sigma(t,T,f(t,T)) = [\,\sigma_0 + \sigma_1(T-t)\,]\, f(t,T).$$

---

[a]Heath, Jarrow, & Morton (1992); Jarrow (1996). The large positive constant κ prevents explosion in finite time.

[b]Gupta & Subrahmanyam (2001, 2005).

[c]Gupta & Subrahmanyam (2001, 2005).

Local-Volatility HJM Models (continued)

- Exponentially dampened volatility proportional to the short rate:[a]

$$\sigma(t,T) = \sigma f(t,t)\, e^{-a(T-t)}.$$

- The Ritchken-Sankarasubramanian (1995) model:[b]

$$\sigma(t,T) = \sigma(t,t)\, e^{-\int_t^T \kappa(x)\, dx}.$$

  - For example,[c]

$$\sigma(t,t) = \sigma r(t)^{\gamma}.$$

[a] Grant & Vora (1999).

[b] The short rate volatility $\sigma(t,t)$ may depend on the short rate $r(t)$.

[c] Ritchken & Sankarasubramanian (1995); Li, Ritchken, & Sankarasubramanian (1995).

Local-Volatility HJM Models (concluded)

- A model attributed to Ian Cooper (1993):[a]

$$\sigma_{\mathrm{p}}(t, T, P(t, T)) = \psi(t) \ln P(t, T)$$

in Eq. (172) on p. 1234:

[a] Rebonato (1996). It is equivalent to the proportional volatility model (173) when $\psi(t)$ is a constant.

Trees for HJM Models

- Obtain today's forward rate curve:

$$f(0,0), f(0,\Delta t), f(0,2\Delta t), f(0,3\Delta t), \ldots, f(0,T).$$

- For binomial trees, generate the two forward rate curves at time Δt:

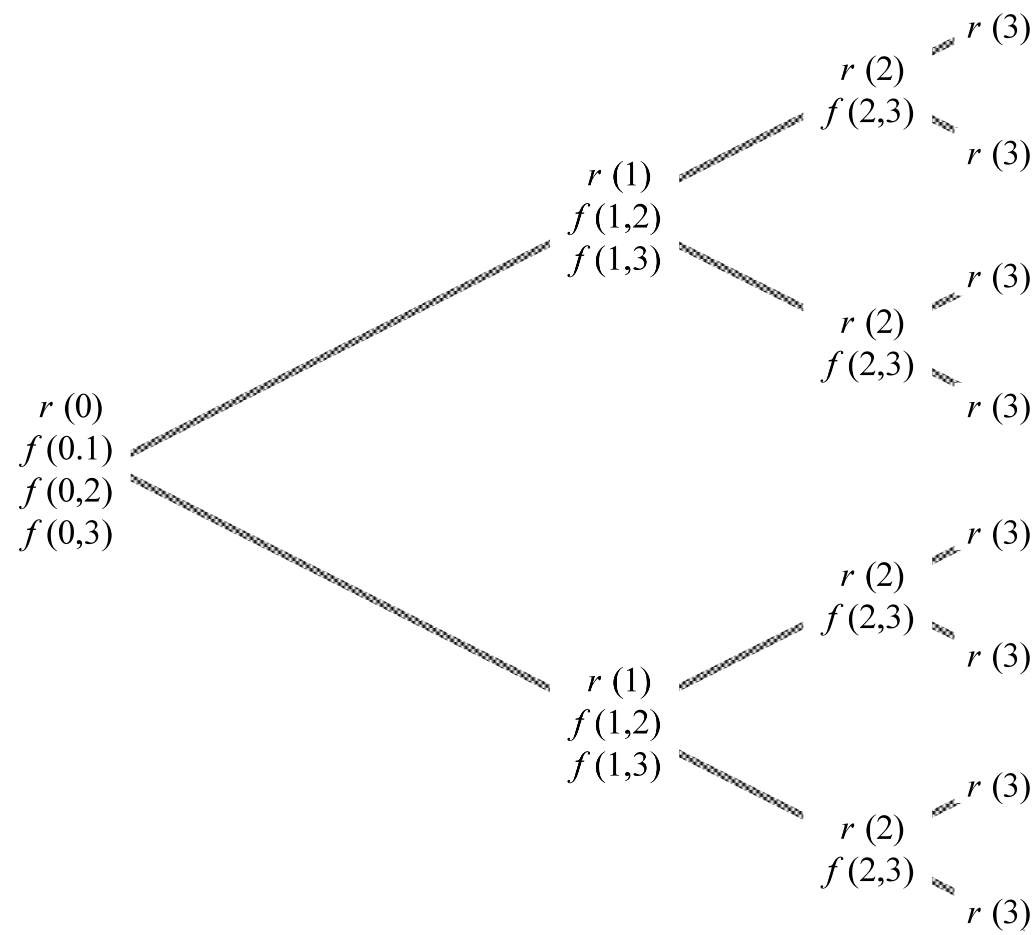
$$f_{\mathrm{u}}(\Delta t,\Delta t), f_{\mathrm{u}}(\Delta t,2\Delta t), f_{\mathrm{u}}(\Delta t,3\Delta t), \ldots, f_{\mathrm{u}}(\Delta t,T),$$
$$f_{\mathrm{d}}(\Delta t,\Delta t), f_{\mathrm{d}}(\Delta t,2\Delta t), f_{\mathrm{d}}(\Delta t,3\Delta t), \ldots, f_{\mathrm{d}}(\Delta t,T).$$

  by Eq. (168) on p. 1226 with $\mu(t,T)$ from Eq. (169) on p. 1228.

Trees for HJM Models (continued)

- Iterate until the maturity $t \leq T$ of the derivative.
- A straightforward implementation of the HJM model results in noncombining trees.
  - For a binomial tree with n time steps, $O(2^n)$ nodes for one-factor HJM models; $O(3^n)$ or $O(4^n)$ for two-factor models.[a]

[a]Clewlow & Strickland (1998); Hull (1999); Nawalkha, Beliaeva, & Soto (2007).



r (0)
f (0.1)
f (0,2)
f (0,3)
r (1)
f (1,2)
f (1,3)
r (1)
f (1,2)
f (1,3)
r (2)
f (2,3)
r (2)
f (2,3)
r (2)
f (2,3)
r (2)
f (2,3)
r (3)
r (3)
r (3)
r (3)
r (3)
r (3)
r (3)
r (3)

Trees for HJM Models (continued)

- Jarrow (1996): "a large number of time steps is not always essential for obtaining good approximations."
- Rebonato (1996): "it is difficult to see how a five-year cap with quarterly resets (let alone an option thereon) could be priced using [10 or 12 time steps]."
- Some trees are not analyzed.[a]

[a] Brace (1996); Gątarek & Kołakowski (2003); Ferris (2012).

Trees for HJM Models (concluded)

- Nawalkha & J. Zhang (2004) has a combining tree for the proportional volatility model with a positive lower bound.
  - It is described in Nawalkha, Beliaeva, & Soto (2007) but not published.
- For Gaussian HJM models, $O(n^2)$ nodes may suffice.[a]

[a] Lok (D99922028), Lu (D00922011), & Lyuu (2020); Lyuu (2019).

Finis