A General Method for Constructing Binomial Models\textsuperscript{a}

- We are given a continuous-time process,
  \[
dy = \alpha(y, t) \, dt + \sigma(y, t) \, dW.
  \]
- Need to make sure the binomial model’s drift and diffusion converge to the above process.
- Set the probability of an up move to
  \[
  \frac{\alpha(y, t) \, \Delta t + y - y_d}{y_u - y_d}.
  \]
- Here \( y_u \triangleq y + \sigma(y, t) \sqrt{\Delta t} \) and \( y_d \triangleq y - \sigma(y, t) \sqrt{\Delta t} \)
  represent the two rates that follow the current rate \( y \).

\textsuperscript{a}Nelson & Ramaswamy (1990).
A General Method (continued)

• The displacements are identical, at $\sigma(y, t) \sqrt{\Delta t}$.

• But the binomial tree may not combine as

$$
\sigma(y, t) \sqrt{\Delta t} - \sigma(y_{u}, t + \Delta t) \sqrt{\Delta t} \\
\neq -\sigma(y, t) \sqrt{\Delta t} + \sigma(y_{d}, t + \Delta t) \sqrt{\Delta t}
$$

in general.

• When $\sigma(y, t)$ is a constant independent of $y$, equality holds and the tree combines.
A General Method (continued)

• To achieve this, define the transformation

\[ x(y, t) \triangleq \int_{y}^{y} \sigma(z, t)^{-1} \, dz. \]

• Then \( x \) follows

\[ dx = m(y, t) \, dt + dW \]

for some \( m(y, t) \).

• The diffusion term is now a constant, and the binomial tree for \( x \) combines.

\[ ^{a}\text{See Exercise 25.2.13 of the textbook.} \]
A General Method (concluded)

• The transformation is unique.\(^a\)

• The probability of an up move remains

\[
\frac{\alpha(y(x, t), t) \Delta t + y(x, t) - y_d(x, t)}{y_u(x, t) - y_d(x, t)},
\]

where \(y(x, t)\) is the inverse transformation of \(x(y, t)\) from \(x\) back to \(y\).

• Note that

\[
\begin{align*}
y_u(x, t) & \triangleq y(x + \sqrt{\Delta t}, t + \Delta t), \\
y_d(x, t) & \triangleq y(x - \sqrt{\Delta t}, t + \Delta t).
\end{align*}
\]

\(^a\)H. Chiu (R98723059) (2012).
Examples

- The transformation is
  \[ \int_{r}^{r} (\sigma \sqrt{z})^{-1} \, dz = \frac{2\sqrt{r}}{\sigma} \]
  for the CIR model.

- The transformation is
  \[ \int_{S}^{S} (\sigma z)^{-1} \, dz = \frac{\ln S}{\sigma} \]
  for the Black-Scholes model \( dS = \mu S \, dt + \sigma S \, dW \).

- The familiar BOPM and CRR discretize \( \ln S \) not \( S \).
On One-Factor Short Rate Models

- By using only the short rate, they ignore other rates on the yield curve.
- Such models also restrict the volatility to be a function of interest rate levels only.
- The prices of all bonds move in the same direction at the same time (their magnitudes may differ).
- The returns on all bonds thus become highly correlated.
- In reality, there seems to be a certain amount of independence between short- and long-term rates.
On One-Factor Short Rate Models (continued)

- One-factor models therefore cannot accommodate nondegenerate correlation structures across maturities.
- Derivatives whose values depend on the correlation structure will be mispriced.
- The calibrated models may not generate term structures as concave as the data suggest.
- The term structure empirically changes in slope and curvature as well as makes parallel moves.
- This is inconsistent with the restriction that all segments of the term structure be perfectly correlated.
On One-Factor Short Rate Models (concluded)

- Multifactor models lead to families of yield curves that can take a greater variety of shapes and can better represent reality.

- But they are much harder to think about and work with.

- They also take much more computer time—the curse of dimensionality.

- These practical concerns limit the use of multifactor models to two- or three-factor ones.\(^a\)

Options on Coupon Bonds

• Assume the market discount function $P$ is a monotonically decreasing function of the short rate $r$.
  - Such as a one-factor short rate model.

• The price of a European option on a coupon bond can be calculated from those on zero-coupon bonds.

• Consider a European call expiring at time $T$ on a bond with par value $\$1$.

• Let $X$ denote the strike price.

\textsuperscript{a}Jamshidian (1989).
Options on Coupon Bonds (continued)

• The bond has cash flows $c_1, c_2, \ldots, c_n$ at times $t_1, t_2, \ldots, t_n$, where $t_i > T$ for all $i$.

• The payoff for the option is

$$\max \left\{ \left[ \sum_{i=1}^{n} c_i P(r(T), T, t_i) \right] - X, 0 \right\}.$$ 

• At time $T$, there is a unique value $r^*$ for $r(T)$ that renders the coupon bond’s price equal the strike price $X$. 
Options on Coupon Bonds (continued)

• This $r^*$ can be obtained by solving

$$X = \sum_{i=1}^{n} c_i P(r, T, t_i)$$

numerically for $r$.

• Let

$$X_i \triangleq P(r^*, T, t_i),$$

the value at time $T$ of a zero-coupon bond with par value $1$ and maturing at time $t_i$ if $r(T) = r^*$.

• Note that $P(r, T, t_i) \geq X_i$ if and only if $r \leq r^*$. 
Options on Coupon Bonds (concluded)

• As \( X = \sum c_i X_i \), the option’s payoff equals

\[
\max \left\{ \left[ \sum_{i=1}^{n} c_i P(r(T), T, t_i) \right] - \left[ \sum_{i=1}^{n} c_i X_i \right], 0 \right\}
\]

\[
= \sum_{i=1}^{n} c_i \times \max(P(r(T), T, t_i) - X_i, 0).
\]

• Thus the call is a package of \( n \) options on the underlying zero-coupon bond.

• Why can’t we do the same thing for Asian options?\(^a\)

\(^a\)Contributed by Mr. Yang, Jui-Chung (D97723002) on May 20, 2009.
No-Arbitrage Term Structure Models
How much of the structure of our theories really tells us about things in nature, and how much do we contribute ourselves?

— Arthur Eddington (1882–1944)

How can I apply this model if I don’t understand it?

— Edward I. Altman (2019)
Motivations

- Recall the difficulties facing equilibrium models mentioned earlier.
  - They usually require the estimation of the market price of risk.\footnote{Recall p. 1099.}
  - They cannot fit the market term structure.
  - But consistency with the market is often mandatory in practice.
No-Arbitrage Models

- No-arbitrage models utilize the full information of the term structure.
- They accept the observed term structure as consistent with an unobserved and unspecified equilibrium.
- From there, arbitrage-free movements of interest rates or bond prices over time are modeled.
- By definition, the market price of risk must be reflected in the current term structure; hence the resulting interest rate process is risk-neutral.
No-Arbitrage Models (concluded)

- No-arbitrage models can specify the dynamics of zero-coupon bond prices, forward rates, or the short rate.
- Bond price and forward rate models are usually non-Markovian (path dependent).
- In contrast, short rate models are generally constructed to be explicitly Markovian (path independent).
- Markovian models are easier to handle computationally.
The Ho-Lee Model

- The short rates at any given time are evenly spaced.
- Let $p$ denote the risk-neutral probability that the short rate makes an up move.
- We shall adopt continuous compounding.

\(^a\)T. Ho & S. B. Lee (1986).
The Ho-Lee Model (continued)

- The Ho-Lee model starts with zero-coupon bond prices $P(t, t + 1), P(t, t + 2), \ldots$ at time $t$ identified with the root of the tree.
- Let the discount factors in the next period be
  \begin{align*}
  P_d(t + 1, t + 2), P_d(t + 1, t + 3), & \ldots, \quad \text{if short rate moves down}, \\
  P_u(t + 1, t + 2), P_u(t + 1, t + 3), & \ldots, \quad \text{if short rate moves up}.
  \end{align*}
- By backward induction, it is not hard to see that for $n \geq 2$,\(^a\)
  \begin{equation}
  P_u(t + 1, t + n) = P_d(t + 1, t + n) e^{-(v_2 + \cdots + v_n)}.
  \end{equation}

\(^a\)See p. 376 of the textbook.
The Ho-Lee Model (continued)

• It is also not hard to check that the \( n \)-period zero-coupon bond has yields

\[
y_d(n) \triangleq -\frac{\ln P_d(t + 1, t + n)}{n - 1} \quad y_u(n) \triangleq -\frac{\ln P_u(t + 1, t + n)}{n - 1} = y_d(n) + \frac{v_2 + \cdots + v_n}{n - 1}
\]

• The volatility of the yield to maturity for this bond is therefore

\[
\kappa_n \triangleq \sqrt{p y_u(n)^2 + (1 - p) y_d(n)^2 - [p y_u(n) + (1 - p) y_d(n)]^2}
\]
\[
= \sqrt{p(1 - p)} (y_u(n) - y_d(n))
\]
\[
= \sqrt{p(1 - p)} \frac{v_2 + \cdots + v_n}{n - 1}.
\]
The Ho-Lee Model (concluded)

• In particular, the short rate volatility is determined by taking $n = 2$:

$$\sigma = \sqrt{p(1 - p)} \; v_2. \quad (156)$$

• The volatility of the short rate therefore equals

$$\sqrt{p(1 - p)} \; (r_u - r_d),$$

where $r_u$ and $r_d$ are the two successor rates.\(^a\)

\(^a\)Contrast this with the lognormal model (133) on p. 1018.
The Ho-Lee Model: Volatility Term Structure

• The volatility term structure is composed of

\[ \kappa_2, \kappa_3, \ldots \]

– The volatility structure is supplied by the market.
– For the Ho-Lee model, it is independent of

\[ r_2, r_3, \ldots \]

• It is easy to compute the \( v_i \)s from the volatility structure, and vice versa.\(^a\)

• The \( r_i \)s can be computed by forward induction.

\(^a\)Review p. 1158.
The Ho-Lee Model: Bond Price Process

• In a risk-neutral economy, the initial discount factors satisfy

\[ P(t, t+n) = [p P_u(t+1, t+n) + (1-p) P_d(t+1, t+n)] P(t, t+1). \]

• Combine the above with Eq. (155) on p. 1157 and assume \( p = 1/2 \) to obtain

\[ P_d(t + 1, t + n) = \frac{P(t, t + n)}{P(t, t + 1)} \frac{2 \times \exp[v_2 + \cdots + v_n]}{1 + \exp[v_2 + \cdots + v_n]}, \]

\[ P_u(t + 1, t + n) = \frac{P(t, t + n)}{P(t, t + 1)} \frac{2}{1 + \exp[v_2 + \cdots + v_n]}. \]

---

\(^{a}\)Recall Eq. (147) on p. 1087.

\(^{b}\)In the limit, only the volatility matters; the first formula is similar to multiple logistic regression.
The Ho-Lee Model: Bond Price Process (concluded)

- The bond price tree combines.\(^a\)

- Suppose all \( v_i \) equal some constant \( v \) and \( \delta \Delta e^v > 0 \).

- Then

\[
P_d(t + 1, t + n) = \frac{P(t, t + n)}{P(t, t + 1)} \frac{2\delta^{n-1}}{1 + \delta^{n-1}},
\]

\[
P_u(t + 1, t + n) = \frac{P(t, t + n)}{P(t, t + 1)} \frac{2}{1 + \delta^{n-1}}.
\]

- Short rate volatility \( \sigma = v/2 \) by Eq. (156) on p. 1159.

- Price derivatives by taking expectations under the risk-neutral probability.

\(^a\)See Exercise 26.2.3 of the textbook.
Calibration

- The Ho-Lee model can be calibrated in $O(n^2)$ time using state prices.
- But it can actually be calibrated in $O(n)$ time.\(^a\)
  - Derive the $v_i$’s in linear time.
  - Derive the $r_i$’s in linear time.

\(^a\)See Programming Assignment 26.2.6 of the textbook.
The Ho-Lee Model: Yields and Their Covariances

- The one-period rate of return of an $n$-period zero-coupon bond is\(^a\)

$$r(t, t + n) \triangleq \ln \left( \frac{P(t + 1, t + n)}{P(t, t + n)} \right).$$

- Its two possible values are

$$\ln \frac{P_d(t + 1, t + n)}{P(t, t + n)} \quad \text{and} \quad \ln \frac{P_u(t + 1, t + n)}{P(t, t + n)}.$$

- Thus the variance of return is\(^b\)

$$\text{Var}[r(t, t + n)] = p(1 - p) [(n - 1) \nu]^2 = (n - 1)^2 \sigma^2.$$

\(^a\)So $r(t, t + n)$ does not mean the $n$-period spot rate at time $t$ here.

\(^b\)Recall that $\sigma$ is the short rate volatility by Eq. (156) on p. 1159.
The Ho-Lee Model: Yields and Their Covariances (concluded)

- The covariance between \( r(t, t + n) \) and \( r(t, t + m) \) is\(^a\)
  \[
  (n - 1)(m - 1)\sigma^2.
  \]

- As a result, the correlation between any two one-period rates of return is one.

- Strong correlation between rates is inherent in all one-factor Markovian models.

\(^a\)See Exercise 26.2.7 of the textbook.
The Ho-Lee Model: Short Rate Process

- The continuous-time limit of the Ho-Lee model is\(^a\)
  \[ dr = \theta(t) \, dt + \sigma \, dW. \quad (157) \]

- This is Vasicek’s model with the mean-reverting drift replaced by a deterministic, time-dependent drift.

- A nonflat term structure of volatilities can be achieved if the short rate volatility is also made time varying,\(^a\)
  \[ dr = \theta(t) \, dt + \sigma(t) \, dW. \]

- This corresponds to the discrete-time model in which \( v_i \) are not all identical.

\(^a\)See Exercise 26.2.10 of the textbook.
The Ho-Lee Model: Some Problems

- Future (nominal) interest rates may be negative.
- The short rate volatility is independent of the rate level.
- It has all the problems associated with a one-factor model.\(^a\)

Problems with No-Arbitrage Models in General

- Interest rate movements should reflect shifts in the model’s state variables (factors) not its parameters.
- Model parameters, such as the drift $\theta(t)$ in the continuous-time Ho-Lee model, should be stable over time.
- But in practice, no-arbitrage models capture yield curve shifts through the recalibration of parameters.
  - A new model is thus born every day.
Problems with No-Arbitrage Models in General (concluded)

• This in effect says the model estimated at some time does not describe the term structure of interest rates and their volatilities at other times.

• Consequently, a model’s intertemporal behavior is suspect, and using it for hedging and risk management may be unreliable.
The Black-Derman-Toy Model\textsuperscript{a}

- This model is extensively used by practitioners.
- The BDT short rate process is the lognormal binomial interest rate process described on pp. 1014ff.\textsuperscript{b}
- The volatility structure\textsuperscript{c} is given by the market.
- From it, the short rate volatilities (thus $v_i$) are determined together with the baseline rates $r_i$.

\textsuperscript{a}Black, Derman, & Toy (BDT) (1990), but essentially finished in 1986 according to Mehrling (2005).
\textsuperscript{b}Repeated on next page.
\textsuperscript{c}Recall Eq. (139) on p. 1065.
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 $\kappa_i$.
- $P_u$ 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,

\[ y_u \triangleq P_u^{-1/(i-1)} - 1, \]
\[ y_d \triangleq P_d^{-1/(i-1)} - 1. \]

• The yield volatility is defined as\(^a\)

\[ \kappa_i \triangleq \frac{\ln(y_u/y_d)}{2}. \]

\(^a\)Recall Eq. (139) on p. 1065.
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,

\[
\frac{P_u + P_d}{2(1 + r_i)} = \frac{1}{(1 + y)^i}. \tag{158}
\]

• 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 \(\kappa_i^2\) denoting their variance, we have

\[
\kappa_i = \frac{1}{2} \ln \left( \frac{P_u^{-1/(i-1)} - 1}{P_d^{-1/(i-1)} - 1} \right). \tag{159}
\]

\(^a\)Recall p. 1174.
The BDT Model: Calibration (continued)

- Solving Eqs. (158)–(159) 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. 1042(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
  $$P_1, P_2, \ldots, P_i.$$  
- They correspond to rates
  $$r, rv, \ldots, rv^{i-1}$$
  for period $i$, respectively.
- One dollar at time $i$ has a present value of
  $$f(r,v) \triangleq \frac{P_1}{1+r} + \frac{P_2}{1+rv} + \frac{P_3}{1+rv^2} + \cdots + \frac{P_i}{1+rv^{i-1}}.$$
The BDT Model: Calibration (continued)

• By Eq. (159) on p. 1177, the yield volatility is

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

• Above, \(P_{u,1}, P_{u,2}, \ldots\) denote the state prices at time \(i - 1\) of the subtree rooted at the up node.\(^a\)

• And \(P_{d,1}, P_{d,2}, \ldots\) denote the state prices at time \(i - 1\) of the subtree rooted at the down node.\(^b\)

\(^a\)Like \(r_2v_2\) on p. 1171.

\(^b\)Like \(r_2\) on p. 1171.
The BDT Model: Calibration (concluded)

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

- Now solve

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

  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)\textsuperscript{a}

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

75MHz Sun SPARCstation 20, one period per year.

\textsuperscript{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.

The Black-Karasinski Model\textsuperscript{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.

\textsuperscript{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

\[
\begin{align*}
    t_2 & \triangleq t_1 + \Delta t_1, \\
    t_3 & \triangleq t_2 + \Delta t_2.
\end{align*}
\]
\[ \ln r_d(t_2) \]
\[ \ln r(t_1) \]
\[ \ln r_u(t_2) \]
\[ \ln r_{du}(t_3) = \ln r_{ud}(t_3) \]
The Black-Karasinski Model: Discrete Time (continued)

• Note that
\[ \ln r_d(t_2) = \ln r(t_1) + \kappa(t_1)(\theta(t_1) - \ln r(t_1)) \Delta t_1 - \sigma(t_1) \sqrt{\Delta t_1}, \]
\[ \ln r_u(t_2) = \ln r(t_1) + \kappa(t_1)(\theta(t_1) - \ln r(t_1)) \Delta t_1 + \sigma(t_1) \sqrt{\Delta t_1}. \]

• To make sure an up move followed by a down move coincides with a down move followed by an up move,
\[
\begin{align*}
\ln r_d(t_2) + \kappa(t_2)(\theta(t_2) - \ln r_d(t_2)) \Delta t_2 + \sigma(t_2) \sqrt{\Delta t_2}, \\
= \ln r_u(t_2) + \kappa(t_2)(\theta(t_2) - \ln r_u(t_2)) \Delta t_2 - \sigma(t_2) \sqrt{\Delta t_2}.
\end{align*}
\]
The Black-Karasinski Model: Discrete Time (continued)

- They imply

\[ \kappa(t_2) = \frac{1 - (\sigma(t_2)/\sigma(t_1)) \sqrt{\Delta t_2/\Delta t_1}}{\Delta t_2}. \]  

(160)

- So from \( \Delta t_1 \), we can calculate the \( \Delta t_2 \) that satisfies the combining condition and then iterate.

\[ t_0 \rightarrow \Delta t_1 \rightarrow t_1 \rightarrow \Delta t_2 \rightarrow t_2 \rightarrow \Delta t_3 \rightarrow \cdots \rightarrow T \]

(roughly).\(^a\)

\(^a\)As \( \kappa(t), \theta(t), \sigma(t) \) are independent of \( r \), the \( \Delta t_i \) will not depend on \( r \) either.
The Black-Karasinski Model: Discrete Time (concluded)

- Unequal durations $\Delta t_i$ 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.

---
aHull & White (1993).
The Extended Vasicek Model\textsuperscript{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.

\textsuperscript{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,
\[ dr = (\theta(t) - ar) \, dt + \sigma \, dW. \]

(161)

• When the current term structure is matched,\(^a\)
\[ \theta(t) = \frac{\partial f(0, t)}{\partial t} + af(0, t) + \frac{\sigma^2}{2a} (1 - e^{-2at}). \]

– 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\textsuperscript{a}

- We describe a trinomial forward induction scheme to calibrate the Hull-White model given \( a \) and \( \sigma \).

- 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 \( j \).

\textsuperscript{a}Hull & White (1993).
The Hull-White Model: Calibration (continued)

- Time increments on the tree are also equally spaced at $\Delta t$ 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

  $t_i \triangleq i\Delta t,
  r_j \triangleq r_0 + j\Delta r,$

  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

\[ \mu_{i,j} \overset{\Delta}{=} \theta(t_i) - ar_j \]  

(162)

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. 1199.

• The middle branch may be an increase of \(\Delta r\), no change, or a decrease of \(\Delta r\).

\(^a\)Or, the annualized expected change.
The Hull-White Model: Calibration (continued)

\[
\begin{align*}
(i, j) & \rightarrow (i + 1, j) \\
& \quad \rightarrow (i + 1, j + 1) \\
& \quad \rightarrow (i + 1, j + 2) \\
(i + 1, j + 1) & \rightarrow (i + 1, j + 2) \\
(i + 1, j + 1) & \rightarrow (i + 1, j + 1) \\
(i + 1, j) & \rightarrow (i + 1, j) \\
& \quad \rightarrow (i + 1, j - 1) \\
& \quad \rightarrow (i + 1, j - 2)
\end{align*}
\]
The Hull-White Model: Calibration (continued)

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

- Define

  \[ p_1(i, j) \triangleq \text{the probability of following the upper branch from node } (i, j), \]
  \[ p_2(i, j) \triangleq \text{the probability of following the middle branch from node } (i, j), \]
  \[ p_3(i, j) \triangleq \text{the probability of following the lower branch from node } (i, j). \]

- 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) \( \theta(t_i) \).
The Hull-White Model: Calibration (continued)

- Once $\theta(t_i)$ is available, $\mu_{i,j}$ can be derived via Eq. (162) on p. 1198.

- 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 $r(0, t_{i+2})$.\(^{a}\)

\(^{a}\text{Not } r(0, t_{i+1})!\)
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 \(\sigma\).

• 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 in turn 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. 1204 for a possible geometry.

• The resulting tree combines.
The probabilities for moving along these branches are functions of $\mu_{i,j}$, $\sigma$, $j$, and $k$:

\[
p_1(i, j) = \frac{\sigma^2 \Delta t + \eta^2}{2(\Delta r)^2} + \frac{\eta}{2\Delta r},
\]

\[
p_2(i, j) = 1 - \frac{\sigma^2 \Delta t + \eta^2}{(\Delta r)^2},
\]

\[
p_3(i, j) = \frac{\sigma^2 \Delta t + \eta^2}{2(\Delta r)^2} - \frac{\eta}{2\Delta r},
\]

where

\[
\eta \triangleq \mu_{i,j} \Delta t + (j - k) \Delta r.
\]
The Hull-White Model: Calibration (continued)

- As trinomial tree algorithms are but explicit methods in disguise,\(^a\) certain relations must hold for \(\Delta r\) and \(\Delta t\) 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, \(\Delta r\) can be set to \(\sigma \sqrt{3\Delta t}\).\(^b\)

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

\(^a\)Recall p. 838.
\(^b\)Hull & White (1988).
The Hull-White Model: Calibration (continued)

- At this point at time $t_i$,

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

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

\[
e^{-r(0,t_{i+2})(i+2)\Delta t} = \sum_j Q(i,j) e^{-r_j \Delta t} \mathbb{E}^{\pi} \left[ e^{-\hat{r}(i+1) \Delta t} \middle| \hat{r}(i) = r_j \right]. \tag{164}
\]

- 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. (164) can be approximated by\(^a\)

\[
E^\pi \left[ e^{-\hat{r}(i+1)\Delta t} \bigg| \hat{r}(i) = r_j \right] \\
\approx e^{-r_j \Delta t} \left( 1 - \mu_{i,j}(\Delta t)^2 + \frac{\sigma^2(\Delta t)^3}{2} \right). \quad (165)
\]

- This solves the chicken-egg problem!

- Substitute Eq. (165) into Eq. (164) and replace \(\mu_{i,j}\) with \(\theta(t_i) - ar_j\) to obtain

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

\(^a\)See Exercise 26.4.2 of the textbook.
The Hull-White Model: Calibration (continued)

• For the Hull-White model, the expectation in Eq. (165) is actually known analytically by Eq. (29) on p. 180:

\[
E^\pi \left[ e^{-\hat{r}(i+1)\Delta t} \bigg| \hat{r}(i) = r_j \right] \\
= e^{-r_j \Delta t + (-\theta(t_i) + ar_j + \sigma^2 \Delta t/2)(\Delta t)^2}.
\]

Therefore, alternatively,

\[
\theta(t_i) = \frac{r(0, t_{i+2})(i + 2)}{\Delta t} + \frac{\sigma^2 \Delta t}{2} + \frac{\ln \sum Q(i, j) e^{-2r_j \Delta t + ar_j (\Delta t)^2}}{(\Delta t)^2}.
\]

• With \(\theta(t_i)\) in hand, we can compute \(\mu_{i,j}\).\(^a\)

\(^a\)See Eq. (162) on p. 1198.
The Hull-White Model: Calibration (concluded)

• With $\mu_{i,j}$ available, we compute the probabilities.\textsuperscript{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?).

\textsuperscript{a}See Eqs. (163) on p. 1205.
Comments on the Hull-White Model

• One can try different values of $a$ and $\sigma$ for each option.

• Or have an $a$ value common to all options but use a different $\sigma$ 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 $\sigma$ 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 $\sigma r^b$.
- But it has at least two shortcomings.
- First, the resulting trinomial tree is irregular (p. 1204).
  - 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. (164) on p. 1208 that helps derive $\theta(t_i)$.

- The resulting $\theta(t_i)$ hence might not yield a tree that matches the spot rates exactly.