Equilibrium Term Structure Models
8. What’s your problem? Any moron can understand bond pricing models.

— Top Ten Lies Finance Professors Tell Their Students
Introduction

• We now survey equilibrium models.

• Recall that the spot rates satisfy

\[ r(t, T) = -\frac{\ln P(t, T)}{T - t} \]

by Eq. (129) on p. 1005.

• Hence the discount function \( P(t, T) \) suffices to establish the spot rate curve.

• All models to follow are short rate models.

• Unless stated otherwise, the processes are risk-neutral.
The Vasicek Model\textsuperscript{a}

- The short rate follows

\[ dr = \beta (\mu - r) \, dt + \sigma \, dW. \]

- The short rate is pulled to the long-term mean level \( \mu \) at rate \( \beta \).

- Superimposed on this “pull” is a normally distributed stochastic term \( \sigma \, dW \).

- Since the process is an Ornstein-Uhlenbeck process,

\[
E[r(T) \mid r(t) = r] = \mu + (r - \mu) e^{-\beta (T - t)}
\]

from Eq. (78) on p. 585.

\textsuperscript{a}Vasicek (1977). Vasicek co-founded KMV, which was sold to Moody’s for USD$210 million in 2002.
The Vasicek Model (continued)

- The price of a zero-coupon bond paying one dollar at maturity can be shown to be

\[ P(t, T) = A(t, T) e^{-B(t,T) r(t)}, \quad (142) \]

where

\[
A(t, T) = \begin{cases} 
\exp \left[ \frac{(B(t,T)-T+t)(\beta^2 \mu - \sigma^2/2)}{\beta^2} - \frac{\sigma^2 B(t,T)^2}{4\beta} \right] & \text{if } \beta \neq 0, \\
\exp \left[ \frac{\sigma^2 (T-t)^3}{6} \right] & \text{if } \beta = 0.
\end{cases}
\]

and

\[
B(t, T) = \begin{cases} 
\frac{1-e^{-\beta(T-t)}}{\beta} & \text{if } \beta \neq 0, \\
T - t & \text{if } \beta = 0.
\end{cases}
\]
The Vasicek Model (concluded)

- If $\beta = 0$, then $P$ goes to infinity as $T \to \infty$.
- Sensibly, $P$ goes to zero as $T \to \infty$ if $\beta \neq 0$.
- Even if $\beta \neq 0$, $P$ may exceed one for a finite $T$.
- The spot rate volatility structure is the curve
  
  \[
  (\partial r(t,T)/\partial r) \sigma = \sigma B(t,T)/(T-t).
  \]

- When $\beta > 0$, the curve tends to decline with maturity.
- The speed of mean reversion, $\beta$, controls the shape of the curve.
- Indeed, higher $\beta$ leads to greater attenuation of volatility with maturity.
A graph showing yield plotted against term. The yield is represented on the Y-axis, ranging from 0.05 to 0.2. The X-axis represents the term, ranging from 2 to 10. Three curves are shown:

- **Normal**: A curve that starts at a higher yield, decreases, and then plateaus.
- **Humped**: A curve that initially decreases, reaches a minimum, and then increases.
- **Inverted**: A curve that starts at a lower yield, increases, and then decreases.

These curves illustrate different yield behaviors over time.
The Vasicek Model: Options on Zeros\textsuperscript{a}

- Consider a European call with strike price $X$ expiring at time $T$ on a zero-coupon bond with par value $\$1$ and maturing at time $s > T$.

- Its price is given by

$$P(t, s) N(x) - XP(t, T) N(x - \sigma_v).$$

\textsuperscript{a}Jamshidian (1989).
The Vasicek Model: Options on Zeros (concluded)

• Above

\[ x \equiv \frac{1}{\sigma_v} \ln \left( \frac{P(t, s)}{P(t, T) X} \right) + \frac{\sigma_v}{2}, \]
\[ \sigma_v \equiv v(t, T) B(T, s), \]
\[ v(t, T)^2 \equiv \begin{cases} \frac{\sigma^2 [1 - e^{-2\beta(T-t)}]}{2\beta}, & \text{if } \beta \neq 0 \\ \sigma^2 (T-t), & \text{if } \beta = 0 \end{cases}. \]

• By the put-call parity, the price of a European put is

\[ XP(t, T) N(-x + \sigma_v) - P(t, s) N(-x). \]
Binomial Vasicek

• Consider a binomial model for the short rate in the time interval \([0, T]\) divided into \(n\) identical pieces.

• Let \(\Delta t \overset{\Delta}{=} T/n\) and

\[
p(r) \overset{\Delta}{=} \frac{1}{2} + \frac{\beta(\mu - r) \sqrt{\Delta t}}{2\sigma}.
\]

• The following binomial model converges to the Vasicek model,

\[
r(k + 1) = r(k) + \sigma \sqrt{\Delta t} \xi(k), \quad 0 \leq k < n.
\]

\(^a\)Nelson & Ramaswamy (1990).
Binomial Vasicek (continued)

• Above, $\xi(k) = \pm 1$ with

$$\text{Prob}[\xi(k) = 1] = \begin{cases} p(r(k)), & \text{if } 0 \leq p(r(k)) \leq 1 \\ 0, & \text{if } p(r(k)) < 0, \\ 1, & \text{if } 1 < p(r(k)). \end{cases}$$

• Observe that the probability of an up move, $p$, is a decreasing function of the interest rate $r$.

• This is consistent with mean reversion.
Binomial Vasicek (concluded)

- The rate is the same whether it is the result of an up move followed by a down move or a down move followed by an up move.
- The binomial tree combines.
- The key feature of the model that makes it happen is its constant volatility, $\sigma$. 
The Cox-Ingersoll-Ross Model\textsuperscript{a}

- It is the following square-root short rate model:
  
  \[
  dr = \beta(\mu - r) \, dt + \sigma \sqrt{r} \, dW. \tag{143}
  \]

- The diffusion differs from the Vasicek model by a multiplicative factor \( \sqrt{r} \).

- The parameter \( \beta \) determines the speed of adjustment.

- The short rate can reach zero only if \( 2\beta \mu < \sigma^2 \).

- See text for the bond pricing formula.

\textsuperscript{a}Cox, Ingersoll, & Ross (1985).
Binomial CIR

- We want to approximate the short rate process in the time interval $[0, T]$.
- Divide it into $n$ periods of duration $\Delta t \triangleq T/n$.
- Assume $\mu, \beta \geq 0$.
- A direct discretization of the process is problematic because the resulting binomial tree will not combine.
Binomial CIR (continued)

- Instead, consider the transformed process
  \[ x(r) \overset{\Delta}{=} 2\sqrt{r}/\sigma. \]

- By Ito’s lemma (p. 562),
  \[ dx = m(x) \, dt + dW, \]
  where
  \[ m(x) \overset{\Delta}{=} 2\beta\mu/(\sigma^2 x) - (\beta x/2) - 1/(2x). \]

- This new process has a constant volatility.
- Thus its binomial tree combines.
Binomial CIR (continued)

• Construct the combining tree for \( r \) as follows.

• First, construct a tree for \( x \).

• Then transform each node of the tree into one for \( r \) via the inverse transformation

\[
    r = f(x) \triangleq \frac{x^2 \sigma^2}{4}
\]

(see p. 1054).

• But when \( x \approx 0 \) (so \( r \approx 0 \)), the moments may not be matched well.\(^a\)

\(^a\)Nawalkha & Beliaeva (2007).
Binomial CIR (continued)

• The probability of an up move at each node $r$ is

$$p(r) \triangleq \frac{\beta(\mu - r) \Delta t + r - r^-}{r^+ - r^-}.$$ 

- $r^+ \triangleq f(x + \sqrt{\Delta t})$ denotes the result of an up move from $r$.
- $r^- \triangleq f(x - \sqrt{\Delta t})$ the result of a down move.

• Finally, set the probability $p(r)$ to one as $r$ goes to zero to make the probability stay between zero and one.
Binomial CIR (concluded)

• It can be shown that

\[
p(r) = \left( \beta \mu - \frac{\sigma^2}{4} \right) \sqrt{\frac{\Delta t}{r}} - B \sqrt{r \Delta t} + C,
\]

for some \( B \geq 0 \) and \( C > 0 \).\(^a\)

• If \( \beta \mu - (\sigma^2/4) \geq 0 \), the up-move probability \( p(r) \) decreases if and only if short rate \( r \) increases.

• Even if \( \beta \mu - (\sigma^2/4) < 0 \), \( p(r) \) tends to decrease as \( r \) increases and decrease as \( r \) declines.

• This phenomenon agrees with mean reversion.

\(^a\)Thanks to a lively class discussion on May 28, 2014.
Numerical Examples

• Consider the process,

\[ 0.2 (0.04 - r) \, dt + 0.1 \sqrt{r} \, dW, \]

for the time interval \([0, 1]\) given the initial rate \(r(0) = 0.04\).

• We shall use \(\Delta t = 0.2\) (year) for the binomial approximation.

• See p. 1058(a) for the resulting binomial short rate tree with the up-move probabilities in parentheses.
Numerical Examples (concluded)

• Consider the node which is the result of an up move from the root.

• Since the root has \( x = 2\sqrt{r(0)/\sigma} = 4 \), this particular node’s \( x \) value equals \( 4 + \sqrt{\Delta t} = 4.4472135955 \).

• Use the inverse transformation to obtain the short rate

\[
\frac{x^2 \times (0.1)^2}{4} \approx 0.0494442719102.
\]

• Once the short rates are in place, computing the probabilities is easy.

• Convergence is quite good (see p. 369 of the textbook).
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

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

• Then \( x \) follows

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

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

• 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

\[
y_u(x, t) \overset{\Delta}{=} y(x + \sqrt{\Delta t}, t + \Delta t),
\]
\[
y_d(x, t) \overset{\Delta}{=} y(x - \sqrt{\Delta t}, t + \Delta t).
\]

\(^a\)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.

• The familiar BOPM and CRR in fact 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)

- Multi-factor 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.
Options on Coupon Bonds\(^a\)

- 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.

\(^a\)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 $X = \sum_{i} c_i X_i$, 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.
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)
Motivations

- Recall the difficulties facing equilibrium models mentioned earlier.
  - They usually require the estimation of the market price of risk.
  - 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.

---

\(^a\)Ho & Lee (1986). Thomas Lee is a “billionaire founder” of Thomas H. Lee Partners LP, according to *Bloomberg* on May 26, 2012.
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\textsuperscript{a}

- 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.

\textsuperscript{a}Ho & 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

\[ P_d(t+1, t+2), P_d(t+1, t+3), \ldots, \quad \text{if short rate moves down}, \]
\[ P_u(t+1, t+2), P_u(t+1, t+3), \ldots, \quad \text{if short rate moves up}. \]

• By backward induction, it is not hard to see that for $n \geq 2$,

\[ P_u(t + 1, t + n) = P_d(t + 1, t + n) e^{-(v_2 + \cdots + v_n)} \]

(144)

(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

$$y_d(n) \triangleq -\frac{\ln P_d(t+1, t+n)}{n-1}$$

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

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

$$\kappa_n \triangleq \sqrt{py_u(n)^2 + (1-p) y_d(n)^2 - [py_u(n) + (1-p) y_d(n)]^2}$$

$$= \sqrt{p(1-p)} \left(y_u(n) - y_d(n)\right)$$

$$= \sqrt{p(1-p)} \frac{v_2 + \cdots + v_n}{n-1}.$$
The Ho-Lee Model (concluded)

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

$$\sigma = \sqrt{p(1 - p)} v_2.$$  \hspace{1cm} (145)

• The variance of the short rate therefore equals

$$p(1 - p)(r_u - r_d)^2,$$

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

\(^a\)Contrast this with the lognormal model (122) on p. 946.
The Ho-Lee Model: Volatility Term Structure

• The volatility term structure is composed of

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

– 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 (review p. 1080).

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

• The volatility structure is supplied by the market.
The Ho-Lee Model: Bond Price Process

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

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

- Combine the above with Eq. (144) on p. 1079 and assume \( p = 1/2 \) to obtain\(^a\)

\[
\begin{align*}
P_d(t + 1, t + n) &= \frac{P(t, t + n)}{P(t, t + 1)} \frac{2 \times \exp[v_2 + \cdots + v_n]}{1 + \exp[v_2 + \cdots + v_n]}, \\
P_u(t + 1, t + n) &= \frac{P(t, t + n)}{P(t, t + 1)} \frac{2}{1 + \exp[v_2 + \cdots + v_n]}. 
\end{align*}
\]

\(^a\)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 \(\delta \equiv e^v > 0\).

• 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. (145) on p. 1081.

• 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.
  – Derive the $v_i$’s in linear time.
  – Derive the $r_i$’s in linear time.$^a$

---

$^a$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

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

• Its two possible value 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

\[
    \text{Var}[r(t, t + n)] = p(1 - p)((n - 1)\nu)^2 = (n - 1)^2\sigma^2.
\]
The Ho-Lee Model: Yields and Their Covariances (concluded)

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

- 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

\[ dr = \theta(t) \, dt + \sigma \, dW. \]

• 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,

\[ dr = \theta(t) \, dt + \sigma(t) \, dW. \]

• 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.
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 everyday.
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. 942ff.\textsuperscript{b}
- The volatility structure is given by the market.
- From it, the short rate volatilities (thus $v_i$) are determined together with $r_i$.

\textsuperscript{a}Black, Derman, & Toy (BDT) (1990), but essentially finished in 1986 according to Mehrling (2005).

\textsuperscript{b}Repeated on next page.
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,

\[
y_{\text{u}} \triangleq P_{\text{u}}^{1/(i-1)} - 1,
\]
\[
y_{\text{d}} \triangleq P_{\text{d}}^{1/(i-1)} - 1.
\]

- The yield volatility is defined as

\[
\kappa_i \triangleq \frac{\ln(y_{\text{u}}/y_{\text{d}})}{2}
\]

(recall Eq. (128) on p. 992).
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 period \(i-1\).
- We now proceed to calculate \(r_i\) and \(v_i\) to extend the tree to period \(i\).
The BDT Model: Calibration (continued)

- 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,

\[
\frac{P_u + P_d}{2(1 + r_1)} = \frac{1}{(1 + y)^i}. \tag{146}
\]

- 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 (p. 1096).

- With \(\kappa_i^2\) denoting their variance, we have
  \[
  \kappa_i = \frac{1}{2} \ln \left( \frac{P_u^{-1/(i-1)} - 1}{P_d^{-1/(i-1)} - 1} \right). \tag{147}
  \]

- Solving Eqs. (146)–(147) for \(r\) and \(v\) with backward induction takes \(O(i^2)\) time.
  - That leads to a cubic-time algorithm.
The BDT Model: Calibration (continued)

- We next employ forward induction to derive a quadratic-time calibration algorithm.\(^a\)

- Recall that forward induction inductively 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. 969(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. (147) on p. 1099, the yield volatility is

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

• Above, \(P_{u,1}, P_{u,2}, \ldots\) denote the state prices at time \(i - 1\) of the subtree rooted at the up node (like \(r_2v_2\) on p. 1093).

• And \(P_{d,1}, P_{d,2}, \ldots\) denote the state prices at time \(i - 1\) of the subtree rooted at the down node (like \(r_2\) on p. 1093).
The BDT Model: Calibration (concluded)

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

- Now solve

\[
\begin{align*}
  f(r, v) &= \frac{1}{(1+y)^i}, \\
  g(r, v) &= \kappa_i,
\end{align*}
\]

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

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

• The short rate volatility clearly should be a declining function of time for the model to display mean reversion.
  – That makes \( \sigma'(t) < 0. \)

• In particular, constant volatility 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

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

• Note that

\[
\begin{align*}
\ln r_d(t_2) &= \ln r(t_1) + \kappa(t_1)(\theta(t_1) - \ln r(t_1)) \Delta t_1 - \sigma(t_1) \sqrt{\Delta t_1}, \\
\ln r_u(t_2) &= \ln r(t_1) + \kappa(t_1)(\theta(t_1) - \ln r(t_1)) \Delta t_1 + \sigma(t_1) \sqrt{\Delta t_1}.
\end{align*}
\]

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

\[ (148) \]

- 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 \( \Delta t_i \)s will not depend on \( r \).
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. 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 different time steps.\textsuperscript{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.

\textsuperscript{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,

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

• When the current term structure is matched,\(^a\)

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

\(^a\)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

- 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 $j$.

---

$^a$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

\[ \mu_{i,j} \overset{\Delta}{=} \theta(t_i) - ar_j \quad (149) \]

\(\mu_{i,j}\) 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. 1121.

- 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)
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) \( \theta(t_i) \).
  
  – 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. (149) on p. 1120.

• 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 \(\sigma\).

- This is done by letting the middle node be as closest to the current short rate \(r_j\) plus the drift \(\mu_{i,j}\Delta t\).\(^b\)

\(^a\)\(p_1(i,j), p_2(i,j), \text{ and } p_3(i,j)\).

\(^b\)A precursor of Lyuu and C. N. 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. 1126 for a possible geometry.
- The resulting tree combines because of the constant jump sizes to reach \( k \) from \( j \).
The Hull-White Model: Calibration (continued)

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

\[
p_1(i, j) = \frac{\sigma^2 \Delta t + \eta^2}{2(\Delta r)^2} + \frac{\eta}{2\Delta r} \tag{150}
\]

\[
p_2(i, j) = 1 - \frac{\sigma^2 \Delta t + \eta^2}{(\Delta r)^2} \tag{150'}
\]

\[
p_3(i, j) = \frac{\sigma^2 \Delta t + \eta^2}{2(\Delta r)^2} - \frac{\eta}{2\Delta r} \tag{150''}
\]

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. 777.
\(^b\)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

$$
e^{-r(0,t_{i+2})(i+2)\Delta t} = \sum_j Q(i, j) e^{-r_j\Delta t} E^\pi \left[ e^{-\hat{r}(i+1)\Delta t} \bigg| \hat{r}(i) = r_j \right]. \quad (151)$$

- The right-hand side represents the value of $1$ obtained by holding a zero-coupon bond until time $t_{i+1}$ and then reinvesting the proceeds at that time at the prevailing short rate $\hat{r}(i + 1)$, which is stochastic.
The Hull-White Model: Calibration (continued)

- The expectation in Eq. (151) can be approximated by

$$E^\pi \left[ e^{-\hat{r}(i+1)\Delta t} \bigg| \hat{r}(i) = r_j \right]$$

$$\approx e^{-r_j\Delta t} \left( 1 - \mu_{i,j}(\Delta t)^2 + \frac{\sigma^2(\Delta t)^3}{2} \right). \quad (152)$$

- This solves the chicken-egg problem!

- Substitute Eq. (152) into Eq. (151) and replace $\mu_{i,j}$ with $\theta(t_i) - ar_j$ to obtain

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

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

• For the Hull-White model, the expectation in Eq. (152) is actually known analytically by Eq. (26) on p. 165:

$$E^\pi \left[ e^{-\hat{r}(i+1)\Delta t} \mid \hat{r}(i) = r_j \right]$$

$$= e^{-r_j \Delta t + (-\theta(t_i) + ar_j + \sigma^2 \Delta t/2)(\Delta t)^2}.$$  

• Therefore, alternatively,

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

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

\(^a\)See Eq. (149) on p. 1120.
The Hull-White Model: Calibration (concluded)

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

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

- 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. (150) on p. 1127.
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. 1126).
  - So it is harder to program (for nonprogrammers).
- The second shortcoming is again a consequence of the tree’s irregular shape.
The Hull-White Model: Calibration with Irregular Trinomial Trees (concluded)

- Recall that the algorithm figured out θ(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. (151) on p. 1130 that helps derive θ(t_i).

- The resulting θ(t_i) hence might not yield a tree that matches the spot rates exactly.
Finis