The Ritchken-Trevor Algorithm (continued)

- The topology of the tree is not a standard combining multinomial tree.
- For example, a few nodes on p. 666 such as nodes $(2, 0)$ and $(2, -1)$ have multiple jump sizes.
- The reason is the path dependence of the model.
  - Two paths can reach node $(2, 0)$ from the root node, each with a different variance for the node.
  - One of the variances results in $\eta = 1$, whereas the other results in $\eta = 2$.

The Ritchken-Trevor Algorithm (concluded)

- The possible values of h_t^2 at a node are exponential nature.
- To address this problem, we record only the maximum and minimum h_t^2 at each node.[a]
- Therefore, each node on the tree contains only two states $(y_t, h_{\max}^2)$ and $(y_t, h_{\min}^2)$.
- Each of $(y_t, h_{\max}^2)$ and $(y_t, h_{\min}^2)$ carries its own η and set of $2n + 1$ branching probabilities.

[a] Cakici and Topyan (2000).

Negative Aspects of the Ritchken-Trevor Algorithm[a]

- A small n may yield inaccurate option prices.
- But the tree will grow exponentially if n is large enough.
  - Specifically, $n > (1 - \beta_1)/\beta_2$ when $r = c = 0$.
- A large n has another serious problem: The tree cannot grow beyond a certain date.
- Thus the choice of n may be limited in practice.
- The RT algorithm can be modified to be free of exponential complexity and shortened maturity.[b]

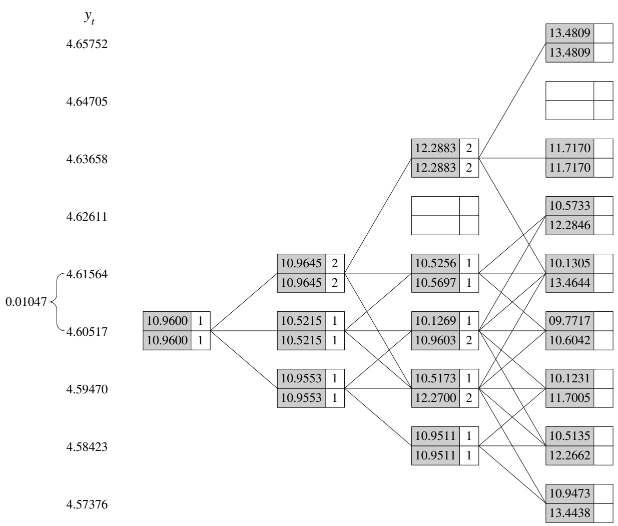
[a] Lyuu and Wu (2003).
[b] Lyuu and Wu (2004).

Numerical Examples

- Assume $S_0 = 100$, $y_0 = \ln S_0 = 4.60517$, $r = 0$, $h_0^2 = 0.0001096$, $\gamma = h_0 = 0.010469$, $n = 1$, $\gamma_n = \gamma/\sqrt{n} = 0.010469$, $\beta_0 = 0.000006575$, $\beta_1 = 0.9$, $\beta_2 = 0.04$, and $c = 0$.
- A daily variance of 0.0001096 corresponds to an annual volatility of $\sqrt{365 \times 0.0001096} \approx 20\%$.
- Let $h^2(i, j)$ denote the variance at node (i, j).
- Initially, $h^2(0, 0) = h_0^2 = 0.0001096$.

Numerical Examples (continued)

- Let $h^2_{\max}(i, j)$ denote the maximum variance at node (i, j).
- Let $h^2_{\min}(i, j)$ denote the minimum variance at node (i, j).
- Initially, $h^2_{\max}(0, 0) = h^2_{\min}(0, 0) = h^2_0$.
- The resulting three-day tree is depicted on p. 672.



A top (bottom) number inside a gray box refers to the minimum (maximum, respectively) variance $h^2_{\min}$ ($h^2_{\max}$, respectively) for the node. Variances are multiplied by 100,000 for readability. A top (bottom) number inside a white box refers to η corresponding to $h^2_{\min}$ ($h^2_{\max}$, respectively).

Numerical Examples (continued)

- Let us see how the numbers are calculated.
- Start with the root node, node $(0, 0)$.
- Try $\eta = 1$ in Eqs. (78)–(80) on p. 657 first to obtain

$$\begin{aligned} p_u &= 0.4974, \\ p_m &= 0, \\ p_d &= 0.5026. \end{aligned}$$

- As they are valid probabilities, the three branches from the root node use single jumps.

Numerical Examples (continued)

- Move on to node $(1, 1)$.
- It has one predecessor node—node $(0, 0)$—and it takes an up move to reach the current node.
- So apply updating rule (81) on p. 663 with $\ell = 1$ and $h_t^2 = h^2(0, 0)$.
- The result is $h^2(1, 1) = 0.000109645$.

Numerical Examples (continued)

- Because $\lfloor h(1, 1)/\gamma \rfloor = 2$, we try $\eta = 2$ in Eqs. (78)–(80) on p. 657 first to obtain

$$\begin{aligned} p_u &= 0.1237, \\ p_m &= 0.7499, \\ p_d &= 0.1264. \end{aligned}$$

- As they are valid probabilities, the three branches from node $(1, 1)$ use double jumps.

Numerical Examples (continued)

- Carry out similar calculations for node $(1, 0)$ with $\ell = 0$ in updating rule (81) on p. 663.
- Carry out similar calculations for node $(1, -1)$ with $\ell = -1$ in updating rule (81).
- Single jump $\eta = 1$ works in both nodes.
- The resulting variances are

$$\begin{aligned} h^2(1, 0) &= 0.000105215, \\ h^2(1, -1) &= 0.000109553. \end{aligned}$$

Numerical Examples (continued)

- Node $(2, 0)$ has 2 predecessor nodes, $(1, 0)$ and $(1, -1)$.
- Both have to be considered in deriving the variances.
- Let us start with node $(1, 0)$.
- Because it takes a middle move to reach the current node, we apply updating rule (81) on p. 663 with $\ell = 0$ and $h_t^2 = h^2(1, 0)$.
- The result is $h_{t+1}^2 = 0.000101269$.

Numerical Examples (continued)

- Now move on to the other predecessor node $(1, -1)$.
- Because it takes an up move to reach the current node, apply updating rule (81) on p. 663 with $\ell = 1$ and $h_t^2 = h^2(1, -1)$.
- The result is $h_{t+1}^2 = 0.000109603$.
- We hence record

$$\begin{aligned} h_{\min}^2(2,0) &= 0.000101269, \\ h_{\max}^2(2,0) &= 0.000109603. \end{aligned}$$

Numerical Examples (continued)

- Consider state $h_{\max}^2(2,0)$ first.
- Because $\lfloor h_{\max}(2,0)/\gamma \rfloor = 2$, we first try $\eta = 2$ in Eqs. (78)–(80) on p. 657 to obtain

$$\begin{aligned} p_u &= 0.1237, \\ p_m &= 0.7500, \\ p_d &= 0.1263. \end{aligned}$$

- As they are valid probabilities, the three branches from node $(2, 0)$ with the maximum variance use double jumps.

Numerical Examples (continued)

- Now consider state $h_{\min}^2(2,0)$.
- Because $\lfloor h_{\min}(2,0)/\gamma \rfloor = 1$, we first try $\eta = 1$ in Eqs. (78)–(80) on p. 657 to obtain

$$\begin{aligned} p_u &= 0.4596, \\ p_m &= 0.0760, \\ p_d &= 0.4644. \end{aligned}$$

- As they are valid probabilities, the three branches from node $(2, 0)$ with the minimum variance use single jumps.

Numerical Examples (continued)

- Node $(2, -1)$ has 3 predecessor nodes.
- Start with node $(1, 1)$.
- Because it takes a down move to reach the current node, we apply updating rule (81) on p. 663 with $\ell = -1$ and $h_t^2 = h^2(1, 1)$.
- The result is $h_{t+1}^2 = 0.0001227$.

Numerical Examples (continued)

- Now move on to predecessor node $(1, 0)$.
- Because it also takes a down move to reach the current node, we apply updating rule (81) on p. 663 with $\ell = -1$ and $h_t^2 = h^2(1, 0)$.
- The result is $h_{t+1}^2 = 0.000105609$.

Numerical Examples (continued)

- Finally, consider predecessor node $(1, -1)$.
- Because it takes a middle move to reach the current node, we apply updating rule (81) on p. 663 with $\ell = 0$ and $h_t^2 = h^2(1, -1)$.
- The result is $h_{t+1}^2 = 0.000105173$.
- We hence record

$$
\begin{aligned}
h_{\min}^2(2, -1) &= 0.000105173, \\
h_{\max}^2(2, -1) &= 0.0001227.
\end{aligned}
$$

Numerical Examples (continued)

- Consider state $h_{\max}^2(2, -1)$.
- Because $\lfloor h_{\max}(2, -1)/\gamma \rfloor = 2$, we first try $\eta = 2$ in Eqs. (78)–(80) on p. 657 to obtain

$$
\begin{aligned}
p_u &= 0.1385, \\
p_m &= 0.7201, \\
p_d &= 0.1414.
\end{aligned}
$$

- As they are valid probabilities, the three branches from node $(2, -1)$ with the maximum variance use double jumps.

Numerical Examples (continued)

- Next, consider state $h_{\min}^2(2, -1)$.
- Because $\lfloor h_{\min}(2, -1)/\gamma \rfloor = 1$, we first try $\eta = 1$ in Eqs. (78)–(80) on p. 657 to obtain

$$
\begin{aligned}
p_u &= 0.4773, \\
p_m &= 0.0404, \\
p_d &= 0.4823.
\end{aligned}
$$

- As they are valid probabilities, the three branches from node $(2, -1)$ with the minimum variance use single jumps.

Numerical Examples (concluded)

- Other nodes at dates 2 and 3 can be handled similarly.
- In general, if a node has k predecessor nodes, then $2k$ variances will be calculated using the updating rule.
  - This is because each predecessor node keeps two variance numbers.
- But only the maximum and minimum variances will be kept.

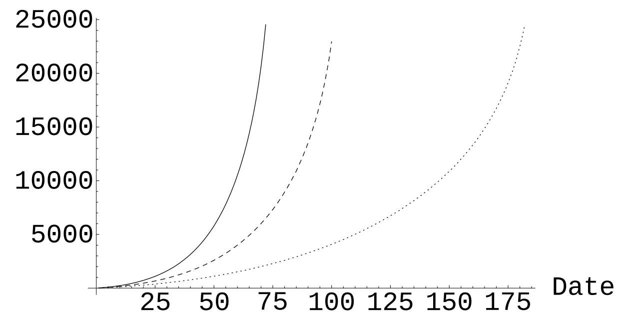
Negative Aspects of the RT Algorithm Revisited[a]

- Recall the problems mentioned on p. 669.
- In our case, combinatorial explosion occurs when

$$n > \frac{1-\beta_1}{\beta_2} = \frac{1-0.9}{0.04} = 2.5.$$

- Suppose we are willing to accept the exponential running time and pick $n = 100$ to seek accuracy.
- But the problem of shortened maturity forces the tree to stop at date 9!

[a]Lyuu and Wu (2003).



Dotted line: $n = 3$; dashed line: $n = 4$; solid line: $n = 5$.

Backward Induction on the RT Tree

- After the RT tree is constructed, it can be used to price options by backward induction.
- Recall that each node keeps two variances $h^2_{\max}$ and $h^2_{\min}$.
- We now increase that number to K equally spaced variances between $h^2_{\max}$ and $h^2_{\min}$ at each node.
- Besides the minimum and maximum variances, the other $K-2$ variances in between are linearly interpolated.

Backward Induction on the RT Tree (continued)

- For example, if $K = 3$, then a variance of 10.5436×10^{-6} will be added between the maximum and minimum variances at node $(2, 0)$ on p. 672.
- In general, the kth variance at node (i, j) is

$$h^2_{\text{min}}(i,j) + k\,\frac{h^2_{\text{max}}(i,j) - h^2_{\text{min}}(i,j)}{K-1},$$

$k = 0, 1, \ldots, K-1$.

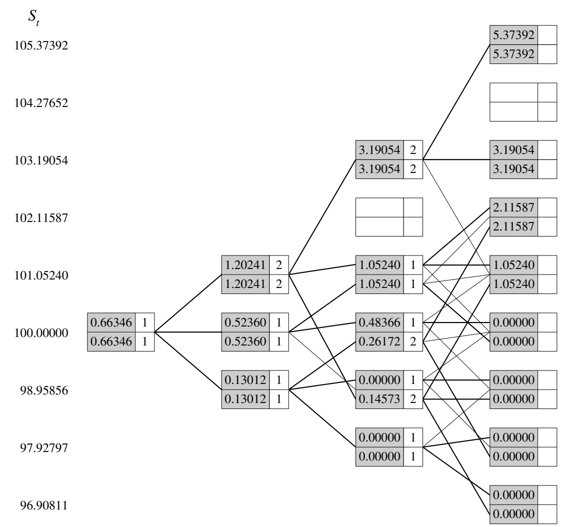
- Each interpolated variance's jump parameter and branching probabilities can be computed as before.

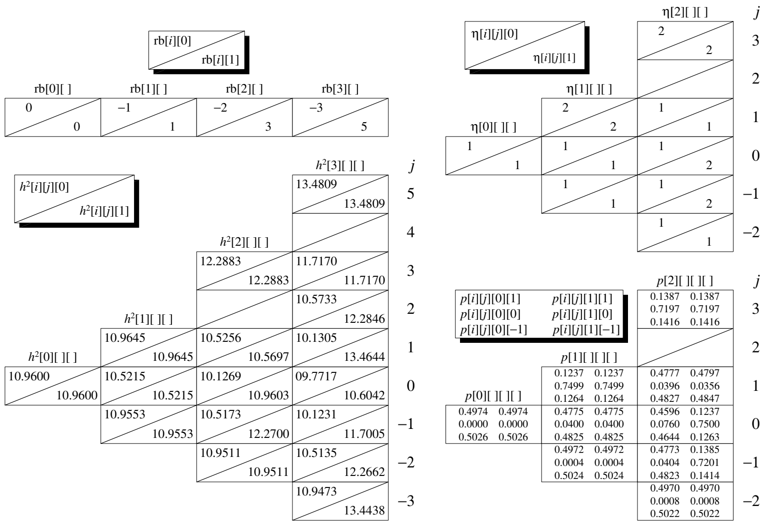
Backward Induction on the RT Tree (concluded)

- During backward induction, if a variance falls between two of the K variances, linear interpolation of the option prices corresponding to the two bracketing variances will be used as the approximate option price.
- The above ideas are reminiscent of the ones on p. 341, where we dealt with arithmetic average-rate options.

Numerical Examples

- We next use the numerical example on p. 672 to price a European call option with a strike price of 100 and expiring at date 3.
- Recall that the riskless interest rate is zero.
- Assume $K = 2$; hence there are no interpolated variances.
- The pricing tree is shown on p. 694 with a call price of 0.66346.
  - The branching probabilities needed in backward induction can be found on p. 695.





Numerical Examples (continued)

- Let us derive some of the numbers on p. 694.
- The option price for a terminal node at date 3 equals $\max(S_3 - 100, 0)$, independent of the variance level.
- Now move on to nodes at date 2.
- The option price at node $(2, 3)$ depends on those at nodes $(3, 5)$, $(3, 3)$, and $(3, 1)$.
- It therefore equals

$$0.1387 \times 5.37392 + 0.7197 \times 3.19054 + 0.1416 \times 1.05240 = 3.19054.$$

- Option prices for other nodes at date 2 can be computed similarly.

Numerical Examples (continued)

- For node $(1, 1)$, the option price for both variances is

$$0.1237 \times 3.19054 + 0.7499 \times 1.05240 + 0.1264 \times 0.14573 = 1.20241.$$

- Node $(1, 0)$ is most interesting.
- We knew that a down move from it gives a variance of 0.000105609.
- This number falls between the minimum variance 0.000105173 and the maximum variance 0.0001227 at node $(2, -1)$ on p. 672.

Numerical Examples (continued)

- The option price corresponding to the minimum variance is 0.
- The option price corresponding to the maximum variance is 0.14573.
- The equation

$$x \times 0.000105173 + (1 - x) \times 0.0001227 = 0.000105609$$

is satisfied by $x = 0.9751$.

- So the option for the down state is approximated by

$$x \times 0 + (1 - x) \times 0.14573 = 0.00362.$$

Numerical Examples (continued)

- The up move leads to the state with option price 1.05240.
- The middle move leads to the state with option price 0.48366.
- The option price at node $(1, 0)$ is finally calculated as

$$0.4775 \times 1.05240 + 0.0400 \times 0.48366 + 0.4825 \times 0.00362 = 0.52360.$$

Numerical Examples (concluded)

- It is possible for some of the three variances following an interpolated variance to exceed the maximum variance or be exceeded by the minimum variance.
- When this happens, the option price corresponding to the maximum or minimum variance will be used during backward induction.
- An interpolated variance may choose a branch that goes into a node that is not reached in the forward-induction tree-building phase.[a]

[a] Lyuu and Wu (2004).

Interest Rate Derivative Securities

Term Structure Fitting

Introduction to Term Structure Modeling

The fox often ran to the hole
by which they had come in,
to find out if his body was still thin enough
to slip through it.
— *Grimm's Fairy Tales*

Outline

- Use the binomial interest rate tree to model stochastic term structure.
  - Illustrates the basic ideas underlying future models.
  - Applications are generic in that pricing and hedging methodologies can be easily adapted to other models.
- Although the idea is similar to the earlier one used in option pricing, the current task is more complicated.
  - The evolution of an entire term structure, not just a single stock price, is to be modeled.
  - Interest rates of various maturities cannot evolve arbitrarily or arbitrage profits may occur.

Issues

- A stochastic interest rate model performs two tasks.
  - Provides a stochastic process that defines future term structures without arbitrage profits.
  - "Consistent" with the observed term structures.
- The unbiased expectations theory, the liquidity preference theory, and the market segmentation theory can all be made consistent with the model.

History

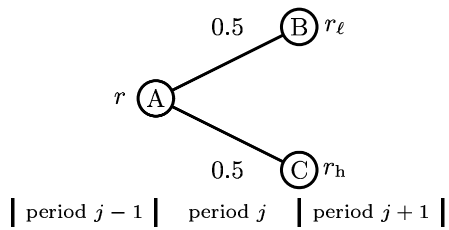
- Methodology founded by Merton (1970).
- Modern interest rate modeling is often traced to 1977 when Vasicek and Cox, Ingersoll, and Ross developed simultaneously their influential models.
- Early models have fitting problems because they may not price today's benchmark bonds correctly.
- An alternative approach pioneered by Ho and Lee (1986) makes fitting the market yield curve mandatory.
- Models based on such a paradigm are called (somewhat misleadingly) arbitrage-free or no-arbitrage models.

Binomial Interest Rate Tree

- Goal is to construct a no-arbitrage interest rate tree consistent with the yields and/or yield volatilities of zero-coupon bonds of all maturities.
  - This procedure is called calibration.
- Pick a binomial tree model in which the logarithm of the future short rate obeys the binomial distribution.
  - Exactly like the CRR tree.
- The limiting distribution of the short rate at any future time is hence lognormal.

Binomial Interest Rate Tree (continued)

- A binomial tree of future short rates is constructed.
- Every short rate is followed by two short rates in the following period (see next page).
- In the figure on p. 710 node A coincides with the start of period j during which the short rate r is in effect.



Binomial Interest Rate Tree (continued)

- At the conclusion of period j, a new short rate goes into effect for period $j+1$.
- This may take one of two possible values:
  - r_ℓ: the "low" short-rate outcome at node B.
  - $r_{\rm h}$: the "high" short-rate outcome at node C.
- Each branch has a fifty percent chance of occurring in a risk-neutral economy.

Binomial Interest Rate Tree (continued)

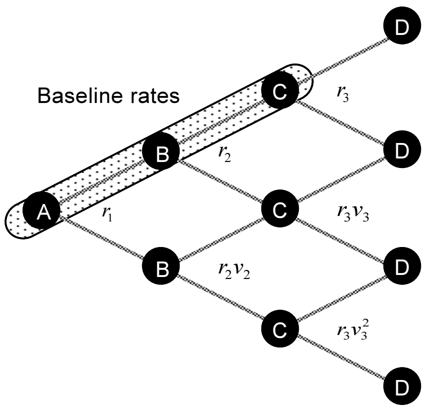
- We shall require that the paths combine as the binomial process unfolds.
- The short rate r can go to $r_{\rm h}$ and r_ℓ with equal risk-neutral probability $1/2$ in a period of length Δt.
- Hence the volatility of $\ln r$ after Δt time is
$$\sigma = \frac{1}{2} \frac{1}{\sqrt{\Delta t}} \ln \left(\frac{r_{\rm h}}{r_\ell} \right)$$
(see Exercise 23.2.3 in text).
- Above, σ is annualized, whereas r_ℓ and $r_{\rm h}$ are period based.

Binomial Interest Rate Tree (continued)

- Note that
$$\frac{r_{\rm h}}{r_\ell} = e^{2\sigma\sqrt{\Delta t}}.$$
- Thus greater volatility, hence uncertainty, leads to larger $r_{\rm h}/r_\ell$ and wider ranges of possible short rates.
- The ratio $r_{\rm h}/r_\ell$ may depend on time if the volatility is a function of time.
- Note that $r_{\rm h}/r_\ell$ has nothing to do with the current short rate r if σ is independent of r.

Binomial Interest Rate Tree (continued)

- In general there are j possible rates in period j,
$$r_j,\, r_j v_j,\, r_j v_j^2,\, \dots,\, r_j v_j^{j-1},$$
where
$$v_j \equiv e^{2\sigma_j \sqrt{\Delta t}} \tag{82}$$
is the multiplicative ratio for the rates in period j (see figure on next page).
- We shall call r_j the baseline rates.
- The subscript j in σ_j is meant to emphasize that the short rate volatility may be time dependent.



Binomial Interest Rate Tree (concluded)

- In the limit, the short rate follows the following process,

$$r(t) = \mu(t)\, e^{\sigma(t)\, W(t)}, \tag{83}$$

  in which the (percent) short rate volatility $\sigma(t)$ is a deterministic function of time.
- As the expected value of $r(t)$ equals $\mu(t)\, e^{\sigma(t)^2 t/2}$, a declining short rate volatility is usually imposed to preclude the short rate from assuming implausibly high values.
- Incidentally, this is how the binomial interest rate tree achieves mean reversion.

Memory Issues

- Path independency: The term structure at any node is independent of the path taken to reach it.
- So only the baseline rates r_i and the multiplicative ratios v_i need to be stored in computer memory.
- This takes up only $O(n)$ space.[a]
- Storing the whole tree would have taken up $O(n^2)$ space.
  - Daily interest rate movements for 30 years require roughly $(30 \times 365)^2/2 \approx 6 \times 10^7$ double-precision floating-point numbers (half a gigabyte!).

[a] Throughout this chapter, n denotes the depth of the tree.

Set Things in Motion

- The abstract process is now in place.
- Now need the annualized rates of return associated with the various riskless bonds that make up the benchmark yield curve and their volatilities.
  - In the U.S., for example, the on-the-run yield curve obtained by the most recently issued Treasury securities may be used as the benchmark curve.

Set Things in Motion (concluded)

- The term structure of (yield) volatilities[a] can be estimated from either the historical data (historical volatility) or interest rate option prices such as cap prices (implied volatility).
- The binomial tree should be consistent with both term structures.
- Here we focus on the term structure of interest rates.

[a] Or simply the volatility (term) structure.

Model Term Structures

- The model price is computed by backward induction.
- Refer back to the figure on p. 710.
- Given that the values at nodes B and C are $P_{\rm B}$ and $P_{\rm C}$, respectively, the value at node A is then
$$\frac{P_{\rm B}+P_{\rm C}}{2(1+r)} + \textsf{cash flow at node A}.$$
- We compute the values column by column without explicitly expanding the binomial interest rate tree (see figure next page).
- This takes quadratic time and linear space.