Financial Computation

Prof. Yuh-Dauh Lyuu
Dept. Computer Science & Information Engineering
and
Department of Finance
National Taiwan University

October/14/2003 by LaTeX

Outline

1. Computational complexity.
2. Derivatives.
3. The binomial model for derivatives pricing.
4. Path-dependent options pricing.
5. Other trees.
6. Monte Carlo pricing.
7. Interest rate models.
8. Model calibration.

References

- Yuh-Dauh Lyuu, *Financial Engineering and Computation: Principles, Mathematics, Algorithms*, Cambridge University Press, 2002.
- Software available at

  `www.csie.ntu.edu.tw/~lyuu/Capitals/capitals.htm`
- Published and unpublished papers.

When Professors Scholes and Merton and I
invested in warrants,
Professor Merton lost the most money.
And I lost the least.
— Fischer Black

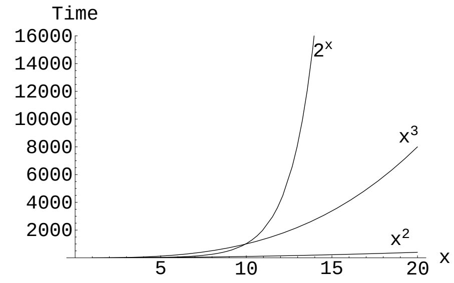
Computational Complexity

It is unworthy of excellent men
to lose hours like slaves
in the labor of computation.
— Gottfried Wilhelm Leibniz (1646–1716)

Measures of Complexity: Time

- Tractable: "Solvable" in *polynomial* time.
  - Such as $O(n)$ and $O(n^2)$.
  - Candidates: Vanilla options.
- Intractable: Otherwise.
  - Candidates: Asian options & certain reset options.
  - Approaches:
    - Analytical approximations.
    - Approximation algorithms.
    - Monte Carlo simulation, etc.

Growths of Various Functions



Measures of Complexity: Space

- Space is usually not an issue.
- It could be an issue for long-dated fixed-income securities or path-dependent derivatives.
- Here is the calculation:
  - A tree has $\approx n^2/2$ nodes, where n is the number of periods until maturity (see later).
  - A 30-year security has more than $365 \times 30 > 10^4$ days.
  - So a tree for a 30-year security has $> 5 \times 10^7$ nodes.

Useful Journals

- *Journal of Computational Finance.*
- *Journal of Derivatives.*
- *Journal of Financial Economics.*
- *Journal of Finance.*
- *Journal of Fixed Income.*
- *Journal of Futures Markets.*
- *Journal of Financial and Quantitative Analysis.*
- *Journal of Real Estate Finance and Economics.*
- *Mathematical Finance.*
- *Review of Financial Studies.*
- *Review of Derivatives Research.*

Introduction

A Very Brief History of Modern Finance

- 1900: Ph.D. thesis *Mathematical Theory of Speculation* of Bachelier (1870–1946).
- 1950s: modern portfolio theory (MPT) of Markowitz.
- 1960s: the Capital Asset Pricing Model (CAPM) of Treynor, Sharpe, Lintner (1916–1984), and Mossin.
- 1960s: the efficient markets hypothesis of Samuelson and Fama.
- 1970s: theory of option pricing of Black (1938–1995) and Scholes.
- 1970s–present: new instruments and pricing methods.

A Very Brief and Biased History of Modern Computers

- 1930s: theory of Gödel (1906–1978), Turing (1912–1954), and Church (1903–1995).
- 1940s: first computers (Z3, ENIAC, etc.) and birth of solid-state transistor (Bell Labs).
- 1950s: Texas Instruments patented integrated circuits; Backus (IBM) invented FORTRAN.
- 1960s: Internet (ARPA) and mainframes (IBM).
- 1970s: relational database (Codd) and PCs (Apple).
- 1980s: IBM PC and Lotus 1-2-3.
- 1990s: Windows 3.1 (Microsoft) and World Wide Web (Berners-Lee).

What This Course Is About

- Financial theories in pricing.
- Mathematical backgrounds.
- Derivative securities.
- Pricing models.
- Efficient algorithms in pricing financial instruments.

Analysis of Algorithms

Computability and Algorithms

- Algorithms are precise procedures that can be turned into computer programs.
- Uncomputable problems.
  - Does this program have infinite loops?
  - Is this program bug free?
- Computable problems.
  - Intractable problems.
  - Tractable problems.

Complexity

- Start with a set of basic operations which will be assumed to take one unit of time.
- The total number of these operations is the total work done by an algorithm (its computational complexity).
- The space complexity is the amount of memory space used by an algorithm.
- Concentrate on the abstract complexity of an algorithm instead of its detailed implementation.
- Complexity is a good guide to an algorithm's *actual* running time.

Asymptotics

- Consider the search algorithm on p. 18.
- The worst-case complexity is n comparisons.
- There are operations besides comparison.
- We care only about the asymptotic growth rate not the exact number of operations.
  - So the complexity of maintaining the loop is subsumed by the complexity of the body of the loop.
- The complexity is hence $O(n)$.

Algorithm for Searching an Element

```
for k = 1, 2, 3, ..., n do
  if x = A[k] then
    return k;
  end if
end for
return not-found;
```


Common Complexities

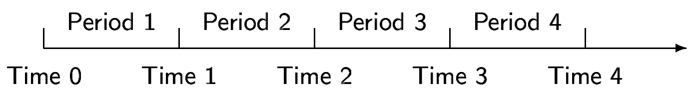
- Let n stand for the "size" of the problem.
  - Number of elements, number of cash flows, etc.
- Linear time if the complexity is $O(n)$.
- Quadratic time if the complexity is $O(n^2)$.
- Cubic time if the complexity is $O(n^3)$.
- Exponential time if the complexity is $2^{O(n)}$.
- Superpolynomial if the complexity is less than exponential but higher than any polynomial.
- It is possible for an exponential-time algorithm to perform well on "typical" inputs.

A Common Misconception about Performance

- A reduction of the running time from 10s to 5s is not as significant as that from 10h to 5h.
- But this is wrong.
  - What if you have 1,000 securities to price.
  - What if you must meet a certain deadline.

Basic Financial Mathematics

The Time Line



Periodic and Continuous Compounding

- If interest is compounded m times per annum,

$$\text{FV} = \text{PV}\left(1 + \frac{r}{m}\right)^{nm}. \tag{1}$$

- As $m \to \infty$ and $(1 + \frac{r}{m})^m \to e^r$ in Eq. (1),

$$\text{FV} = \text{PV}e^{rn},$$

  where $e = 2.71828\ldots$.

- If the annual interest rate is r_1 for n_1 years and r_2 for the following n_2 years, the FV of \$1 will be

$$e^{r_1 n_1 + r_2 n_2}.$$

Efficient Algorithms for PV and FV

- The PV of the cash flow $C_1, C_2, \ldots, C_n$ at times $1, 2, \ldots, n$ is

$$\frac{C_1}{1+y} + \frac{C_2}{(1+y)^2} + \cdots + \frac{C_n}{(1+y)^n}.$$

- Computed by the algorithm on p. 25 in time $O(n)$.

Algorithm for Evaluating PV

```
x := 0;
d := 1 + y;
for i = n, n − 1, . . . , 1 do
    x := (x + C_i)/d;
end for
return x;
```


The Idea Behind p. 25: Horner's Rule

- This idea is

$$\left(\cdots\left(\left(\frac{C_n}{1+y}+C_{n-1}\right)\frac{1}{1+y}+C_{n-2}\right)\frac{1}{1+y}+\cdots\right)\frac{1}{1+y}.$$

  - Due to Horner (1786–1837) in 1819.
- It is the most efficient possible in terms of the absolute number of arithmetic operations.

Yields

- The term yield denotes the return of investment.
- Two widely used yields are the bond equivalent yield (BEY) and the mortgage equivalent yield (MEY).
  - BEY corresponds to the r in Eq. (1) on p. 23 that equates PV with FV when $m = 2$.
  - MEY corresponds to the r in Eq. (1) on p. 23 that equates PV with FV when $m = 12$.

Internal Rate of Return (IRR)

- It is the interest rate which equates an investment's PV with its price P,

$$P=\frac{C_1}{(1+y)}+\frac{C_2}{(1+y)^2}+\frac{C_3}{(1+y)^3}+\cdots+\frac{C_n}{(1+y)^n}.$$

- IRR assumes all cash flows are reinvested at the *same* rate as the internal rate of return.
- The above formula is the foundation upon which pricing methodologies are built.

Numerical Methods for Yields

- Solve $f(y) = 0$ for $y \geq -1$, where

$$f(y) \equiv \sum_{t=1}^{n} \frac{C_t}{(1+y)^t} - P.$$

  - P is the market price.
- The function $f(y)$ is monotonic in y if $C_t > 0$ for all t.
- Hence a unique solution exists.

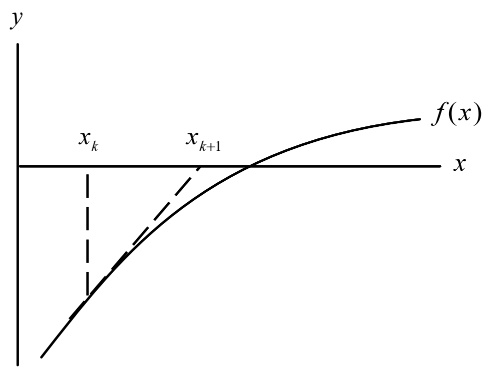
The Newton-Raphson Method

- Start with a first approximation x_0 to a root of $f(x) = 0$.
- Then

$$x_{k+1} \equiv x_k - \frac{f(x_k)}{f'(x_k)}.$$

- When computing yields,

$$f'(x) = -\sum_{t=1}^{n} \frac{tC_t}{(1+x)^{t+1}}.$$



Solving Systems of Nonlinear Equations

- The Newton-Raphson method can be extended to higher dimensions.
- Let (x_k, y_k) be the kth approximation to the solution of the two simultaneous equations,

$$\begin{aligned} f(x,y) &= 0, \\ g(x,y) &= 0. \end{aligned}$$

Solving Systems of Nonlinear Equations (concluded)

- The $(k+1)$st approximation (x_{k+1}, y_{k+1}) satisfies the following linear equations,

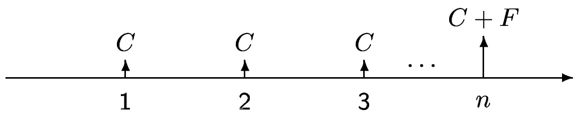
$$\begin{bmatrix} \frac{\partial f(x_k,y_k)}{\partial x} & \frac{\partial f(x_k,y_k)}{\partial y} \\ \frac{\partial g(x_k,y_k)}{\partial x} & \frac{\partial g(x_k,y_k)}{\partial y} \end{bmatrix} \begin{bmatrix} \Delta x_{k+1} \\ \Delta y_{k+1} \end{bmatrix} = - \begin{bmatrix} f(x_k, y_k) \\ g(x_k, y_k) \end{bmatrix},$$

  where $\Delta x_{k+1} \equiv x_{k+1} - x_k$ and $\Delta y_{k+1} \equiv y_{k+1} - y_k$.

- The above has a unique solution for $(\Delta x_{k+1}, \Delta y_{k+1})$ when the 2×2 matrix is invertible.
- Set $(x_{k+1}, y_{k+1}) = (x_k + \Delta x_{k+1}, y_k + \Delta y_{k+1})$.

Level-Coupon Bonds

- Coupon rate.
- Par value, paid at maturity.
- F denotes the par value and C denotes the coupon.
- Cash flow:



Pricing Formula

$$\begin{aligned} P &= \sum_{i=1}^{n} \frac{C}{\left(1+\frac{r}{m}\right)^i} + \frac{F}{\left(1+\frac{r}{m}\right)^n} \\ &= C \, \frac{1-\left(1+\frac{r}{m}\right)^{-n}}{\frac{r}{m}} + \frac{F}{\left(1+\frac{r}{m}\right)^n}. \end{aligned}$$

- n: number of cash flows.
- m: number of payments per year.
- r: annual rate compounded m times per annum.
- $C = Fc/m$ when c is the annual coupon rate.
- Price P can be computed in $O(1)$ time.

Bond Price Volatility

The Key Question: Price Volatility

- Volatility measures how bond prices respond to interest rate changes.
- It is key to the risk management of interest-rate-sensitive securities.
- Assume level-coupon bonds throughout.
- Define price volatility as the sensitivity of the percentage price change to changes in interest rates,

$$-\frac{\frac{\partial P}{\partial y}}{P}.$$

Price Volatility of Bonds

- The price volatility of a coupon bond is

$$-\frac{(C/y)\,n - \left(C/y^2\right)\left((1+y)^{n+1} - (1+y)\right) - nF}{(C/y)\left((1+y)^{n+1} - (1+y)\right) + F(1+y)},$$

  where F is the par value, and C is the coupon payment per period.
- Hence computable in constant time.
- For bonds without embedded options,

$$-\frac{\frac{\partial P}{\partial y}}{P} > 0.$$

Macaulay Duration

- The Macaulay duration (MD) is a weighted average of the times to an asset's cash flows.
- The weights are the cash flows' PVs divided by the asset's price,

$$\text{MD} \equiv \frac{1}{P} \sum_{i=1}^{n} \frac{iC_i}{(1+y)^i}.$$

- The Macaulay duration, in periods, is equal to

$$\text{MD} = -(1+y)\,\frac{\partial P/P}{\partial y}. \tag{2}$$

MD of Bonds

- The MD of a coupon bond is

$$\text{MD} = \frac{1}{P}\left[\sum_{i=1}^{n} \frac{iC}{(1+y)^i} + \frac{nF}{(1+y)^n}\right]. \tag{3}$$

- Can be simplified to

$$\text{MD} = \frac{c(1+y)\,[\,(1+y)^n - 1\,] + ny(y-c)}{cy\,[\,(1+y)^n - 1\,] + y^2},$$

  where c is the period coupon rate.
- The MD of a zero-coupon bond is its term to maturity n.

Finesse

- Equations (2) on p. 39 and (3) on p. 40 hold only if the coupon C, the par value F, and the maturity n are all independent of the yield y.
- That is, if the cash flow is independent of yields.
- To see this point, suppose the market yield declines.
- The MD will be lengthened.
- For securities whose maturity actually decreases as a result, the MD may actually decrease.

How Not To Think of MD

- The MD has its origin in measuring the length of time a bond investment is outstanding.
- But you use it that way at your peril.
- The MD should be seen mainly as measuring price volatility.
- Many, if not most, duration-related terminology cannot be comprehended otherwise.

Modified Duration

- Modified duration is defined as

$$\text{modified duration} \equiv -\frac{\partial P/P}{\partial y} = \frac{\text{MD}}{(1+y)}. \tag{4}$$

- By Taylor expansion,

$$\text{percent price change} \approx -\text{modified duration} \times \text{yield change}.$$

- Both MD and modified duration are easy to calculate.

Effective Duration

- A general numerical formula for volatility is the effective duration,

$$\frac{P_- - P_+}{P_0(y_+ - y_-)}. \tag{5}$$

  - P_- is the price if the yield is decreased by Δy.
  - P_+ is the price if the yield is increased by Δy.
  - P_0 is the initial price, y is the initial yield.
  - Δy is small.
- One can compute the effective duration of just about any financial instrument.

Effective Duration (concluded)

- It is most useful where yield changes alter the cash flow or securities whose cash flow is so complex that simple formulas are unavailable.
- An alternative is to use

$$\frac{P_0 - P_+}{P_0 \, \Delta y}.$$

- This is more economical but less accurate.

Convexity

- Convexity is defined as

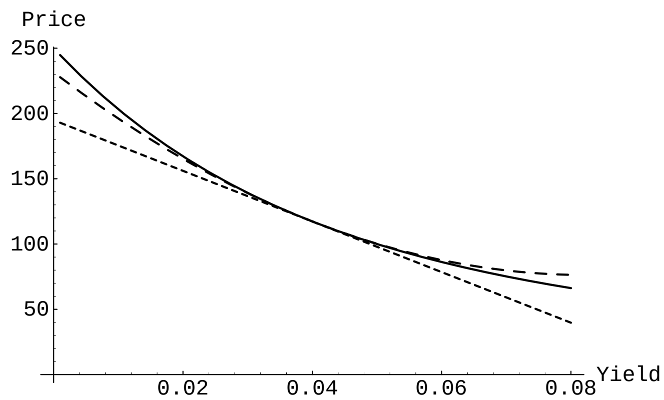
$$\text{convexity (in periods)} \equiv \frac{\partial^2 P}{\partial y^2} \, \frac{1}{P}. \tag{6}$$

- For a bond with positive convexity, the price rises more for a rate decline than it falls for a rate increase of equal magnitude.
- Hence, between two bonds with the same duration, the one with a higher convexity is more valuable.

Use of Convexity

- The approximation $\Delta P / P \approx -$ duration $\times$ yield change works for small yield changes.
- To improve upon it for larger yield changes, use

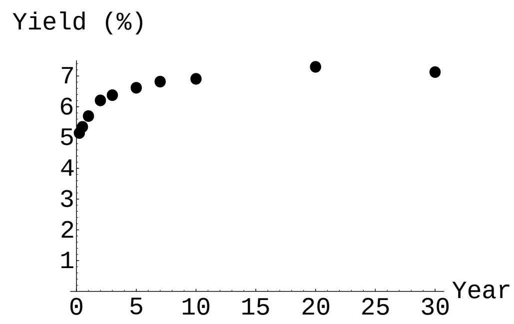
$$\begin{aligned} \frac{\Delta P}{P} &\approx \frac{\partial P}{\partial y} \frac{1}{P} \Delta y + \frac{1}{2} \frac{\partial^2 P}{\partial y^2} \frac{1}{P} (\Delta y)^2 \\ &= -\text{duration} \times \Delta y + \frac{1}{2} \times \text{convexity} \times (\Delta y)^2. \end{aligned}$$



Term Structure of Interest Rates

Term Structure of Interest Rates

- How do interest rates change with maturity?
- The set of yields to maturity for bonds forms the term structure.
  - The bonds must be of equal quality.
  - They differ solely in their terms to maturity.
- The term structure is fundamental to the valuation of fixed-income securities.



Term Structure of Interest Rates (concluded)

- *Term structure* often refers exclusively to the yields of zero-coupon bonds.
- A *yield curve* plots yields to maturity against maturity.
- A *par* yield curve is constructed from bonds trading near par.

Spot Rates

- The i-period spot rate $S(i)$ is the yield to maturity of an i-period zero-coupon bond.
- The PV of one dollar i periods from now is

$$[\,1+S(i)\,]^{-i}.$$

- The one-period spot rate is called the short rate.
- A spot rate curve is a plot of spot rates against maturity.

Spot Rate Discount Methodology

- A level-coupon bond has the price

$$P=\sum_{i=1}^{n}\frac{C}{[\,1+S(i)\,]^{i}}+\frac{F}{[\,1+S(n)\,]^{n}}. \qquad (7)$$

- This pricing method incorporates information from the term structure.
- Discount each cash flow at the corresponding spot rate.

Discount Factors

- In general, any riskless security having a cash flow $C_1, C_2, \ldots, C_n$ should have a market price of

$$P=\sum_{i=1}^{n}C_i d(i).$$

  - Above, $d(i) \equiv [\,1+S(i)\,]^{-i}$, $i=1,2,\ldots,n$, are called discount factors.
  - $d(i)$ is the PV of one dollar i periods from now.
- The discount factors are often interpolated to form a continuous function called the discount function.

Extracting Spot Rates from Yield Curve

- Start with the short rate $S(1)$.
  - Note that short-term Treasuries are zero-coupon bonds.
- Compute $S(2)$ from the two-period coupon bond price P by solving

$$P=\frac{C}{1+S(1)}+\frac{C+100}{[\,1+S(2)\,]^{2}}.$$

Extracting Spot Rates from Yield Curve (continued)

- Inductively, we are given the market price P of the n-period coupon bond and $S(1), S(2), \ldots, S(n-1)$.
- Then $S(n)$ can be computed from Eq. (7), repeated below,

$$P = \sum_{i=1}^{n} \frac{C}{[\,1+S(i)\,]^i} + \frac{F}{[\,1+S(n)\,]^n}.$$

- The total running time is $O(n)$.

Extracting Spot Rates from Yield Curve (concluded)

$S[\,1\,] := (100/P[\,1\,]) - 1;$
$p := P[\,1\,]/100;$
for $i = 2, 3, \ldots, n$ **do**
  Solve $P[\,i\,] = C[\,i\,] \times p + (C[\,i\,] + 100)/(1+x)^i$ for x;
  $S[\,i\,] := x;$
  $p := p + (1+x)^{-i};$
end for
`return` $S[\,]$;

Static Spread

- Consider a *risky* bond with the cash flow $C_1, C_2, \ldots, C_n$ and selling for P.
- Since riskiness must be compensated,

$$P < \sum_{t=1}^{n} \frac{C_t}{[\,1+S(t)\,]^t}.$$

- The static spread is the amount s such that

$$P = \sum_{t=1}^{n} \frac{C_t}{[\,1+s+S(t)\,]^t}.$$

- Can be computed by the Newton-Raphson method.

Fundamental Statistical Concepts

There are three kinds of lies:
lies, damn lies, and statistics.
— Benjamin Disraeli (1804–1881)

One death is a tragedy,
but a million deaths are a statistic.
— Josef Stalin (1879–1953)

Moments

- The variance of a random variable X is defined as
$$\text{Var}[X] \equiv E\left[(X - E[X])^2\right].$$
- The covariance between random variables X and Y is
$$\text{Cov}[X, Y] \equiv E[(X - \mu_X)(Y - \mu_Y)],$$
where μ_X and μ_Y are the means of X and Y, respectively.
- Random variables X and Y are uncorrelated if $\text{Cov}[X, Y] = 0$.

The Normal Distribution

- A random variable X has the normal distribution with mean μ and variance σ^2 if its probability density function is $e^{-(x-\mu)^2/(2\sigma^2)}/(\sigma\sqrt{2\pi})$.
- This is expressed by $X \sim N(\mu, \sigma^2)$.
- The standard normal distribution has zero mean, unit variance, and the distribution function
$$\text{Prob}[X \leq z] = N(z) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2}\, dx.$$

Generation of Univariate Normal Distributions

- Let X be uniformly distributed over $(0, 1]$ so that $\text{Prob}[X \leq x] = x$ for $0 < x \leq 1$.
- Repeatedly draw two samples x_1 and x_2 from X until $\omega \equiv (2x_1 - 1)^2 + (2x_2 - 1)^2 < 1$.
- Then $c(2x_1 - 1)$ and $c(2x_2 - 1)$ are independent standard normal variables where
$$c \equiv \sqrt{-2(\ln \omega)/\omega}\,.$$

A Dirty Trick

- Let ξ_i are independent and uniformly distributed over $(0, 1)$.
- A simple method to generate the standard normal variable is to calculate

$$\sum_{i=1}^{12} \xi_i - 6,$$

- Always blame your random number generator last; instead, check your programs first.

Generation of Bivariate Normal Distributions

- Pairs of normally distributed variables with correlation ρ can be generated.
- Let X_1 and X_2 be independent standard normal variables.
- Then

$$\begin{aligned} U &\equiv aX_1, \\ V &\equiv \rho U + \sqrt{1-\rho^2}\, aX_2, \end{aligned}$$

are the desired random variables with $\text{Var}[U] = \text{Var}[V] = a^2$ and $\text{Cov}[U, V] = \rho a^2$.

The Lognormal Distribution

- A random variable Y is said to have a lognormal distribution if $\ln Y$ has a normal distribution.
- Let $X \sim N(\mu, \sigma^2)$ and $Y \equiv e^X$.
- The mean and variance of Y are

$$\mu_Y = e^{\mu+\sigma^2/2} \text{ and } \sigma_Y^2 = e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right), \quad (8)$$

respectively.

Option Basics

The shift toward options as
the center of gravity of finance [. . .]
— Merton H. Miller (1923–2000)

Calls and Puts

- A call gives its holder the right to buy a number of the underlying asset by paying a strike price.
- A put gives its holder the right to sell a number of the underlying asset for the strike price.
- An embedded option has to be traded along with the underlying asset.
- How to price options?

Exercise

- When a call is exercised, the holder pays the strike price in exchange for the stock.
- When a put is exercised, the holder receives from the writer the strike price in exchange for the stock.
- An option can be exercised prior to the expiration date: early exercise.

American and European

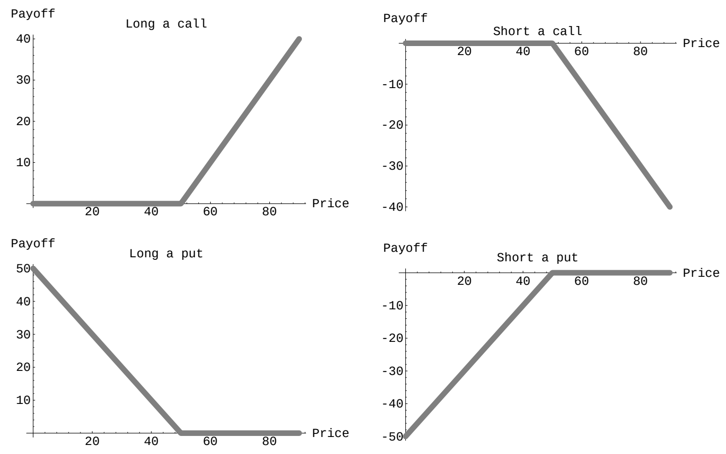
- American options can be exercised at any time up to the expiration date.
- European options can only be exercised at expiration.
- An American option is worth at least as much as an otherwise identical European option because of the early exercise feature.

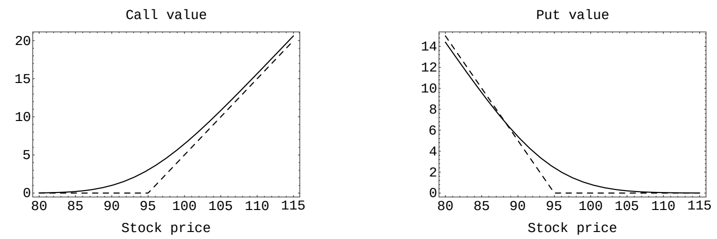
Convenient Conventions

- C: call value.
- P: put value.
- X: strike price.
- S: stock price.
- D: dividend.

Payoff

- The payoff of a call at expiration is $C = \max(0, S - X)$.
- The payoff of a put at expiration is $P = \max(0, X - S)$.
- At any time t before the expiration date, we call $\max(0, S_t - X)$ the intrinsic value of a call.
- At any time t before the expiration date, we call $\max(0, X - S_t)$ the intrinsic value of a put.
- Finding an option's value at any time before expiration is a major intellectual breakthrough.





Cash Dividends

- Exchange-traded stock options are not cash dividend-protected.
  - The option contract is not adjusted for cash dividends.
- The stock price falls by an amount roughly equal to the amount of the cash dividend as it goes ex-dividend.

Arbitrage in Option Pricing

Arbitrage

- The no-arbitrage principle says there should be no free lunch.
- It supplies the argument for option pricing.
- A riskless arbitrage opportunity is one that, without any initial investment, generates nonnegative returns under all circumstances and positive returns under some.
- In an efficient market, such opportunities do not exist.

Relative Option Prices

- These relations hold regardless of the probabilistic model for stock prices.
- Assume, among other things, that there are no transactions costs or margin requirements, borrowing and lending are available at the riskless interest rate, interest rates are nonnegative, and there are no arbitrage opportunities.
- The put-call parity[a]:

$$C = P + S - \text{PV}(X).$$

[a] Castelli (1877).

Option Pricing Models

The Setting

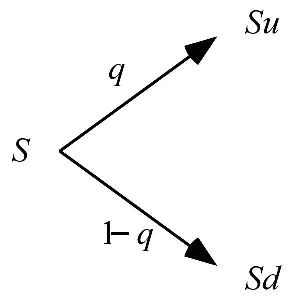
- The no-arbitrage principle is insufficient to pin down the exact option value without further assumptions on the probabilistic behavior of stock prices.
- One major obstacle is that it seems a risk-adjusted interest rate is needed to discount the option's payoff.
- Breakthrough came in 1973 when Black (1938–1995) and Scholes with help from Merton published their celebrated option pricing model.
- Known as the Black-Scholes option pricing model.

Terms and Approach

- C: call value.
- P: put value.
- X: strike price
- S: stock price
- $\hat{r} > 0$: the continuously compounded riskless rate per period.
- $R \equiv e^{\hat{r}}$: gross return.
- Start from the discrete-time binomial model.

Binomial Option Pricing Model (BOPM)

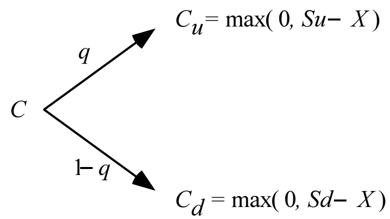
- Time is discrete and measured in periods.
- If the current stock price is S, it can go to Su with probability q and Sd with probability $1 - q$, where $0 < q < 1$ and $d < u$.
  - In fact, $d < R < u$ must hold to rule out arbitrage.
- Six pieces of information suffice to determine the option value based on arbitrage considerations: S, u, d, X, $\hat{r}$, and the number of periods to expiration.



Call on a Non-Dividend-Paying Stock: Single Period

- The expiration date is only one period from now.
- C_u is the price at time one if the stock price moves to Su.
- C_d is the price at time one if the stock price moves to Sd.
- Clearly,

$$\begin{aligned} C_u &= \max(0, Su - X), \\ C_d &= \max(0, Sd - X). \end{aligned}$$



Call Pricing in One Period

- Set up a portfolio of h shares of stock and B dollars in riskless bonds.
  - This costs $hS + B$.
  - We call h the hedge ratio or delta.
- The value of this portfolio at time one is either $hSu + RB$ or $hSd + RB$.
- Choose h and B such that the portfolio replicates the payoff of the call,

$$\begin{aligned} hSu + RB &= C_u, \\ hSd + RB &= C_d. \end{aligned}$$

Call Pricing in One Period (concluded)

- Solve the above equations to obtain

$$h = \frac{C_u - C_d}{Su - Sd} \geq 0, \tag{9}$$

$$B = \frac{uC_d - dC_u}{(u-d)\,R}. \tag{10}$$

- By the no-arbitrage principle, the European call should cost the same as the equivalent portfolio, $C = hS + B$.
- As $uC_d - dC_u < 0$, the equivalent portfolio is a levered long position in stocks.

American Call Pricing in One Period

- Have to consider immediate exercise.
- $C = \max(hS + B, S - X)$.
  - When $hS + B \geq S - X$, the call should not be exercised immediately.
  - When $hS + B < S - X$, the option should be exercised immediately.
- For non-dividend-paying stocks, early exercise is not optimal, so $C = hS + B$.

Put Pricing in One Period

- Puts can be similarly priced.
- The delta for the put is $(P_u - P_d)/(Su - Sd) \leq 0$, where

$$P_u = \max(0, X - Su),$$
$$P_d = \max(0, X - Sd).$$

- Let $B = \frac{uP_d - dP_u}{(u-d)\,R}$.
- The European put is worth $hS + B$.
- The American put is worth $\max(hS + B, X - S)$.

Pseudo Probability

- After substitution and rearrangement,

$$hS + B = \frac{\left(\frac{R-d}{u-d}\right) C_u + \left(\frac{u-R}{u-d}\right) C_d}{R}. \tag{11}$$

- Rewrite Eq. (11) as

$$hS + B = \frac{pC_u + (1-p)\,C_d}{R},$$

where

$$p \equiv \frac{R-d}{u-d}.$$

- As $0 < p < 1$, it may be interpreted as a probability.

Risk-Neutral Probability

- The expected rate of return for the stock is equal to the riskless rate $\hat{r}$ under $q = p$ as $pSu + (1 - p)\,Sd = RS$.
- Risk-neutral investors care only about expected returns.
- The expected rates of return of all securities must be the riskless rate when investors are risk-neutral.
- For this reason, p is called the risk-neutral probability.
- The value of an option is the expectation of its discounted future payoff in a risk-neutral economy.
- So the rate used for discounting the FV is the riskless rate in a risk-neutral economy.

Binomial Distribution

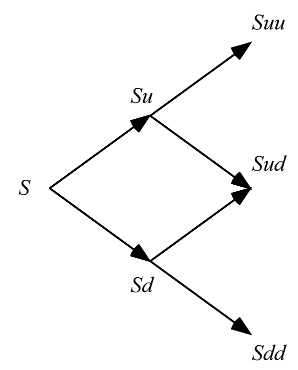
- Denote the binomial distribution with parameters n and p by

$$b(j; n, p) \equiv \binom{n}{j} p^j (1-p)^{n-j} = \frac{n!}{j!\,(n-j)!}\, p^j (1-p)^{n-j}.$$

  - $n! = n \times (n-1) \cdots 2 \times 1$ with the convention $0! = 1$.
- Suppose you toss a coin n times with p being the probability of getting heads.
- Then $b(j; n, p)$ is the probability of getting j heads.

Option on a Non-Dividend-Paying Stock: Multi-Period

- Consider a call with two periods remaining before expiration.
- Under the binomial model, the stock can take on three possible prices at time two: Suu, Sud, and Sdd.
  - Note that the tree combines.
- At any node, the next two stock prices only depend on the current price, not the prices of earlier times.



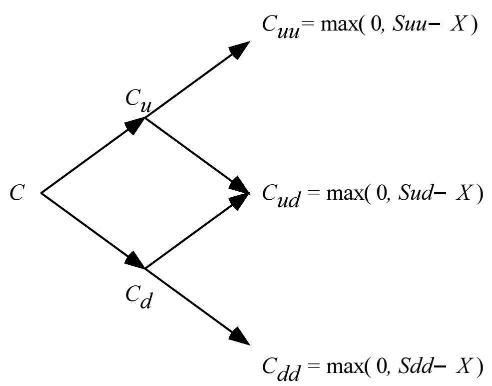
Option on a Non-Dividend-Paying Stock: Multi-Period (continued)

- Let C_{uu} be the call's value at time two if the stock price is Suu.
- Thus,

$$C_{uu} = \max(0, Suu - X).$$

- C_{ud} and C_{dd} can be calculated analogously,

$$\begin{aligned} C_{ud} &= \max(0, Sud - X), \\ C_{dd} &= \max(0, Sdd - X). \end{aligned}$$



Option on a Non-Dividend-Paying Stock: Multi-Period (continued)

- The call values at time one can be obtained by applying the same logic:

$$\begin{aligned} C_u &= \frac{pC_{uu} + (1-p)\,C_{ud}}{R}, \qquad (12) \\ C_d &= \frac{pC_{ud} + (1-p)\,C_{dd}}{R}. \end{aligned}$$

- Deltas can be derived from Eq. (9) on p. 89.
- For example, the delta at C_u is

$$(C_{uu} - C_{ud})/(Suu - Sud).$$

Option on a Non-Dividend-Paying Stock: Multi-Period (concluded)

- We now reach the current period.
- An equivalent portfolio of h shares of stock and $\$B$ riskless bonds can be set up for the call that costs C_u (C_d, resp.) if the stock price goes to Su (Sd, resp.).
- The values of h and B can be derived from Eqs. (9)–(10) on p. 89.

Early Exercise

- Since the call will not be exercised at time one even if it is American, $C_u \geq Su - X$ and $C_d \geq Sd - X$.
- Therefore,

$$\begin{aligned} hS + B &= \frac{pC_u + (1-p)\, C_d}{R} \geq \frac{[\, pu + (1-p)\, d\,]\, S - X}{R} \\ &= S - \frac{X}{R} > S - X. \end{aligned}$$

- So the call again will not be exercised at present, and

$$C = hS + B = \frac{pC_u + (1-p)\, C_d}{R}.$$

Backward Induction (Zermelo)

- The above expression calculates C from the two successor nodes C_u and C_d and none beyond.
- The same computation happens at C_u and C_d, too, as demonstrated in Eq. (12) on p. 99.
- This recursive procedure is called backward induction.
- Now, C equals

$$\begin{aligned} & [\, p^2 C_{uu} + 2p(1-p)\, C_{ud} + (1-p)^2 C_{dd}](1/R^2) \\ = & [\, p^2 \cdot \max\left(0, Su^2 - X\right) + 2p(1-p) \cdot \max\left(0, Sud - X\right) \\ & +(1-p)^2 \cdot \max\left(0, Sd^2 - X\right)](1/R^2). \end{aligned}$$

Backward Induction (continued)

- In the n-period case,

$$C = \frac{\sum_{j=0}^{n} \binom{n}{j} p^j (1-p)^{n-j} \times \max\left(0, Su^j d^{n-j} - X\right)}{R^n}.$$

  - The value of a call on a non-dividend-paying stock is the expected discounted payoff at expiration in a risk-neutral economy.
- The value of a European put is

$$P = \frac{\sum_{j=0}^{n} \binom{n}{j} p^j (1-p)^{n-j} \times \max\left(0, X - Su^j d^{n-j}\right)}{R^n}.$$

Risk-Neutral Pricing Methodology

- Every derivative can be priced as if the economy were risk-neutral.
- For a European-style derivative with the terminal payoff function $\mathcal{D}$, its value is

$$e^{-\hat{r}n} E^{\pi}[\,\mathcal{D}\,].$$

  - E^{π} means the expectation is taken under the risk-neutral probability.

The Binomial Option Pricing Formula

- Let a be the minimum number of upward price moves for the call to finish in the money.
- So a is the smallest nonnegative integer such that $Su^a d^{n-a} \geq X$, or

$$a = \left\lceil \frac{\ln(X/Sd^n)}{\ln(u/d)} \right\rceil.$$

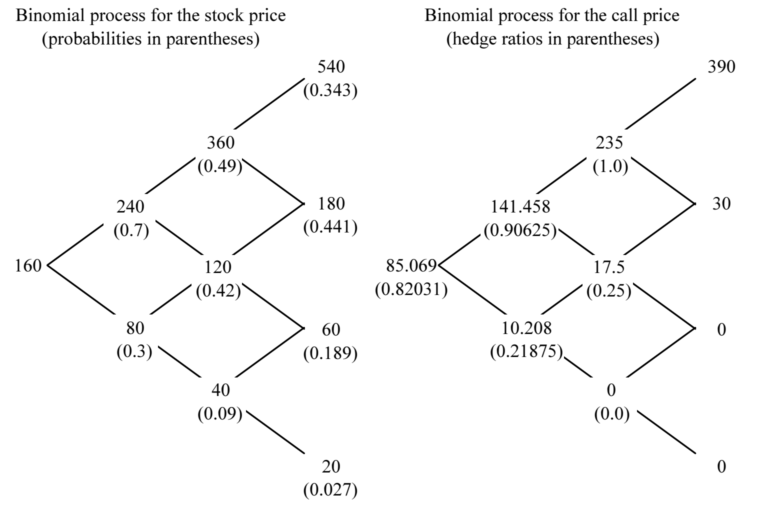
- Hence,

$$C = \frac{\sum_{j=a}^{n} \binom{n}{j} p^j (1-p)^{n-j} \left(Su^j d^{n-j} - X \right)}{R^n}.$$

Numerical Examples

- A non-dividend-paying stock is selling for \$160.
- $u = 1.5$ and $d = 0.5$.
- $r = 18.232\%$ per period.
- Consider a European call on this stock with $X = 150$ and $n = 3$.
- The call value is \$85.069 by backward induction.
- Also the PV of the expected payoff at expiration,

$$\frac{390 \times 0.343 + 30 \times 0.441}{(1.2)^3} = 85.069.$$



Numerical Examples (continued)

- Mispricing leads to arbitrage profits.
- Suppose the option is selling for \$90 instead.
- Sell the call for \$90 and invest \$85.069 in the replicating portfolio with 0.82031 shares of stock required by delta.
- Borrow $0.82031 \times 160 - 85.069 = 46.1806$ dollars.
- The fund that remains, $90 - 85.069 = 4.931$ dollars, is the arbitrage profit as we will see.

Numerical Examples (continued)

Time 1:

- Suppose the stock price moves to \$240.
- The new delta is 0.90625.
- Buy $0.90625 - 0.82031 = 0.08594$ more shares at the cost of $0.08594 \times 240 = 20.6256$ dollars financed by borrowing.
- Debt now totals $20.6256 + 46.1806 \times 1.2 = 76.04232$ dollars.

Numerical Examples (continued)

Time 2:

- Suppose the stock price plunges to \$120.
- The new delta is 0.25.
- Sell $0.90625 - 0.25 = 0.65625$ shares for an income of $0.65625 \times 120 = 78.75$ dollars.
- Use this income to reduce the debt to $76.04232 \times 1.2 - 78.75 = 12.5$ dollars.

Numerical Examples (continued)

Time 3 (the case of rising price):

- The stock price moves to \$180.
- The call we wrote finishes in the money.
- For a loss of $180 - 150 = 30$ dollars, close out the position by either buying back the call or buying a share of stock for delivery.
- Financing this loss with borrowing brings the total debt to $12.5 \times 1.2 + 30 = 45$ dollars.
- It is repaid by selling the 0.25 shares of stock for $0.25 \times 180 = 45$ dollars.

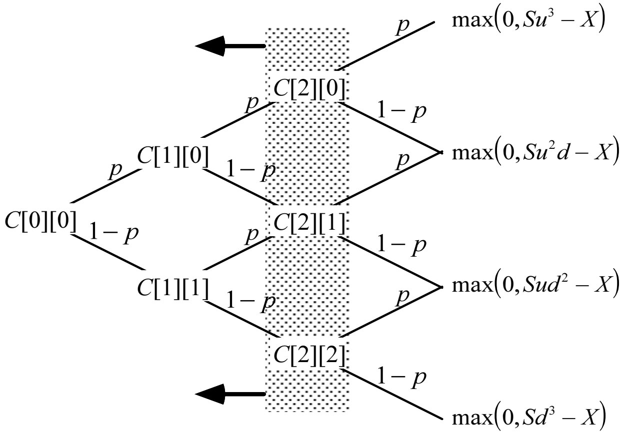
Numerical Examples (concluded)

Time 3 (the case of declining price):

- The stock price moves to \$60.
- The call we wrote is worthless.
- Sell the 0.25 shares of stock for a total of $0.25 \times 60 = 15$ dollars.
- Use it to repay the debt of $12.5 \times 12 = 15$ dollars.

Binomial Tree Algorithms for European Options

- The BOPM implies the binomial tree algorithm that applies backward induction.
- The total running time is $O(n^2)$.
- The memory requirement is $O(n)$.
- To price European puts, simply replace the payoff.



Toward the Black-Scholes Formula

- As the number of periods, n, increases, the stock price ranges over ever larger numbers of possible values, and trading takes place nearly continuously.
- A proper calibration of the model parameters makes the BOPM converge to the continuous-time model.

Toward the Black-Scholes Formula (continued)

- Let τ denote the time to expiration of the option measured in years.
- Let r be the continuously compounded annual rate.
- Pick

$$u = e^{\sigma\sqrt{\tau/n}}, \quad d = e^{-\sigma\sqrt{\tau/n}}, \quad q = \frac{1}{2} + \frac{1}{2}\,\frac{\mu}{\sigma}\sqrt{\frac{\tau}{n}}.$$

- Other choices are possible (see text).
- The risk-neutral probability may lie outside $[0,1]$.
- The problems disappear when $n > r^2\tau/\sigma^2$.

Toward the Black-Scholes Formula (continued)

- As
$$n \to \infty,$$
$\ln S_\tau$ approaches the normal distribution with mean $\mu\tau + \ln S$ and variance $\sigma^2\tau$.
- S_τ has a lognormal distribution in the limit.

Toward the Black-Scholes Formula (concluded)

Theorem 1 (The Black-Scholes Formula)

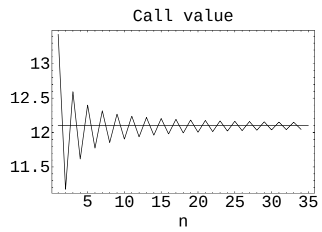
$$\begin{aligned} C &= SN(x) - Xe^{-r\tau} N(x - \sigma\sqrt{\tau}), \\ P &= Xe^{-r\tau} N(-x + \sigma\sqrt{\tau}) - SN(-x), \end{aligned}$$

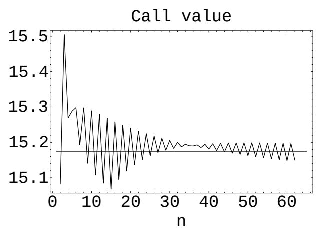
where

$$x \equiv \frac{\ln(S/X) + \left(r + \sigma^2/2\right)\tau}{\sigma\sqrt{\tau}}.$$

BOPM and Black-Scholes Model

- The Black-Scholes formula needs five parameters: S, X, σ, τ, and r.
- Binomial tree algorithms take six inputs: S, X, u, d, $\hat{r}$, and n.
- The connections are
$$u = e^{\sigma\sqrt{\tau/n}}, \; d = e^{-\sigma\sqrt{\tau/n}}, \; \hat{r} = r\tau/n.$$
- The binomial tree algorithms converge reasonably fast.
- Oscillations can be eliminated by the judicious choices of u and d (see text).





Implied Volatility

- Volatility is the sole parameter not directly observable.
- The Black-Scholes formula can be used to compute the market's opinion of the volatility.
  - Solve for σ given the option price, S, X, τ, and r with numerical methods.
  - How about American options?
- This volatility is called the implied volatility.
- Implied volatility is often preferred to historical volatility in practice.

Binomial Tree Algorithms for American Puts

- Early exercise has to be considered.
- The binomial tree algorithm starts with the terminal payoffs $\max(0, X - Su^j d^{n-j})$ and applies backward induction.
- At each intermediate node, it checks for early exercise by comparing the payoff if exercised with continuation.

Options on a Stock That Pays Dividends

- Early exercise must be considered.
- Proportional dividend payout model is tractable (see text).
  - The dividend amount is a constant proportion of the prevailing stock price.
- In general, the corporate dividend policy is a complex issue.

Continuous Dividend Yields

- Dividends are Paid continuously.
  - Approximates a broad-based stock market portfolio.
- The payment of a continuous dividend yield at rate q reduces the growth rate of the stock price by q.

Continuous Dividend Yields (continued)

- The Black-Scholes formulas hold with S replaced by $Se^{-q\tau}$:[a]

$$C = Se^{-q\tau} N(x) - Xe^{-r\tau} N(x - \sigma\sqrt{\tau}), \tag{13}$$

$$P = Xe^{-r\tau} N(-x + \sigma\sqrt{\tau}) - Se^{-q\tau} N(-x), \tag{13$'$}$$

where

$$x \equiv \frac{\ln(S/X) + \left(r - q + \sigma^2/2\right)\tau}{\sigma\sqrt{\tau}}.$$

[a] Merton (1973).

Continuous Dividend Yields (concluded)

- To run binomial tree algorithms, pick the risk-neutral probability as

$$\frac{e^{(r-q)\,\Delta t} - d}{u - d}, \tag{14}$$

where $\Delta t \equiv \tau/n$.

- The u and d remain unchanged.
- Other than the change in Eq. (14), binomial tree algorithms stay the same.

Extensions of Options Theory

Barrier Options

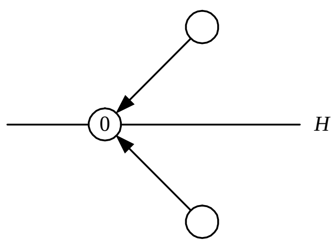
- Their payoff depends on whether the underlying asset's price reaches a certain price level H.
- A knock-out option is an ordinary European option which ceases to exist if the barrier H is reached by the price of its underlying asset.
- A call knock-out option is sometimes called a down-and-out option if $H < S$.
- A put knock-out option is sometimes called an up-and-out option when $H > S$.

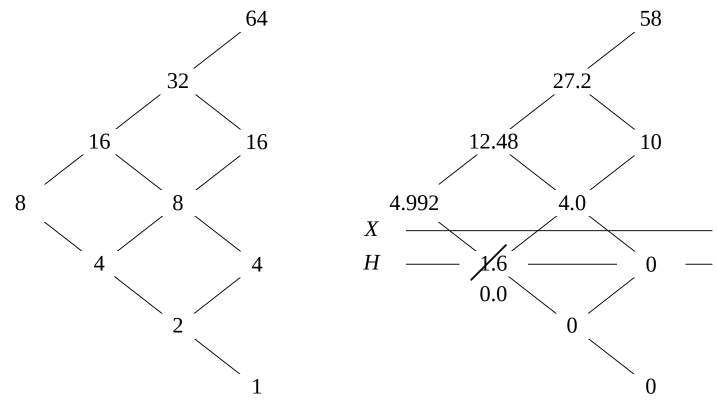
Barrier Options (concluded)

- A knock-in option comes into existence if a certain barrier is reached.
- A down-and-in option is a call knock-in option that comes into existence only when the barrier is reached and $H < S$.
- An up-and-in is a put knock-in option that comes into existence only when the barrier is reached and $H > S$.
- Formulas exist for all kinds of barrier options.

Binomial Tree Algorithms

- Barrier options can be priced by binomial tree algorithms.
- Below is for the down-and-out option.

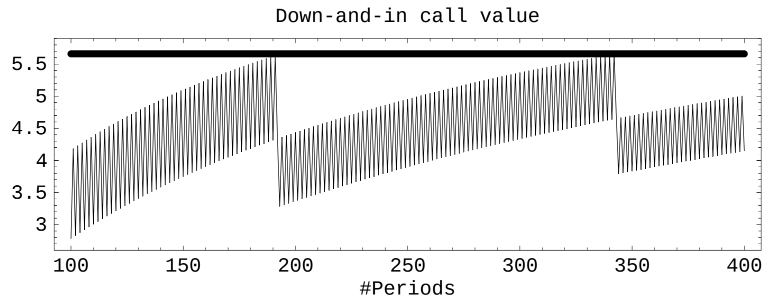




$S = 8$, $X = 6$, $H = 4$, $R = 1.25$, $u = 2$, and $d = 0.5$.
Backward induction: $C = (0.5 \times C_u + 0.5 \times C_d)/1.25$.

Binomial Tree Algorithms (concluded)

- But convergence is erratic because H is not at a price level on the tree (see plot on next page).
  - Typically, the barrier has to be adjusted to be at a price level.
- Hence the algorithms are useless in practice.
- Solutions will be presented later.



Path-Dependent Derivatives

- Let $S_0, S_1, \ldots, S_n$ denote the prices of the underlying asset over the life of the option.
- S_0 is the known price at time zero.
- S_n is the price at expiration.
- The standard European call has a terminal value depending only on the last price, $\max(S_n - X, 0)$.
- Its value thus depends only on the underlying asset's terminal price regardless of how it gets there.

Path-Dependent Derivatives (continued)

- In contrast, some derivatives are path-dependent in that their terminal payoffs depend "critically" on the paths.
- The (arithmetic) average-rate call has a terminal value given by

$$\max\left(\frac{1}{n+1}\sum_{i=0}^{n} S_i - X, 0\right).$$

- The average-rate put's terminal value is given by

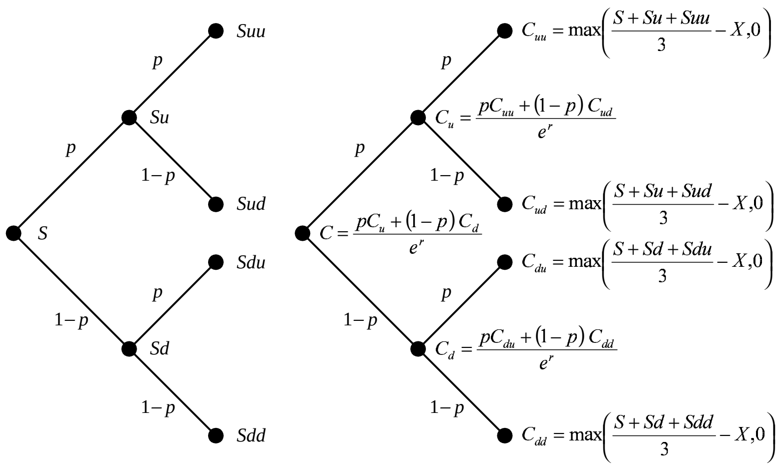
$$\max\left(X - \frac{1}{n+1}\sum_{i=0}^{n} S_i, 0\right).$$

Path-Dependent Derivatives (concluded)

- Average-rate options are also called Asian options.
- They are useful hedging tools for firms that will make a stream of purchases over a time period.
  - The costs are likely to be linked to the average price.

Average-Rate Options

- Average-rate options are notoriously hard to price.
- The binomial tree for the averages does not combine (see p. 138).
- A straightforward algorithm is to enumerate the 2^n price paths for an n-period binomial tree and then average the payoffs.
- But the exponential complexity makes it impractical.
- As a result, the Monte Carlo method and approximation algorithms are some of the alternatives left.



Approximation Algorithm for Asian Options

- Based on the BOPM.
- Consider a node at time j with the underlying asset price equal to $S_0 u^{j-i} d^i$.
- Name such a node $N(j, i)$.
- The running sum $\sum_{m=0}^{j} S_m$ at this node has a maximum value of

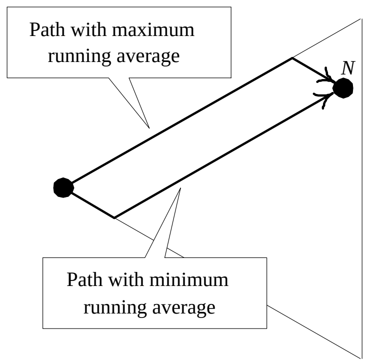
$$S_0(1 + \overbrace{u + u^2 + \cdots + u^{j-i} + u^{j-i}d + \cdots + u^{j-i}d^i}^{j})$$
$$= S_0 \frac{1 - u^{j-i+1}}{1-u} + S_0 u^{j-i} d \, \frac{1-d^i}{1-d}.$$

Approximation Algorithm for Asian Options (continued)

- Divide this value by $j+1$ and call it $A_{\max}(j, i)$.
- Similarly, the running sum has a minimum value of

$$S_0(1 + \overbrace{d + d^2 + \cdots + d^i + d^i u + \cdots + d^i u^{j-i}}^{j})$$
$$= S_0 \frac{1 - d^{i+1}}{1-d} + S_0 d^i u \, \frac{1-u^{j-i}}{1-u}.$$

- Divide this value by $j+1$ and call it $A_{\min}(j, i)$.
- $A_{\min}$ and $A_{\max}$ are running averages.



Approximation Algorithm for Asian Options (continued)

- The possible running averages at $N(j,i)$ are far too many: $\binom{j}{i}$.
- But all lie between $A_{\min}(j,i)$ and $A_{\max}(j,i)$.
- Pick $k+1$ equally spaced values in this range and treat them as the true and only running averages:

$$A_m(j,i) \equiv \left(\frac{k-m}{k}\right) A_{\min}(j,i) + \left(\frac{m}{k}\right) A_{\max}(j,i)$$

for $m = 0, 1, \ldots, k$.

Approximation Algorithm for Asian Options (continued)

- Such "bucketing" introduces some errors.
- Backward induction calculates the option values at each node for these $k+1$ running averages.
- Suppose the current node is $N(j,i)$ and the running average is a.
- Assume the next node is $N(j+1,i)$, after an up move.
- As the asset price there is $S_0 u^{j+1-i} d^i$, we seek the option value corresponding to the running average

$$A_{\rm u} \equiv \frac{(j+1)\,a + S_0 u^{j+1-i} d^i}{j+2}.$$

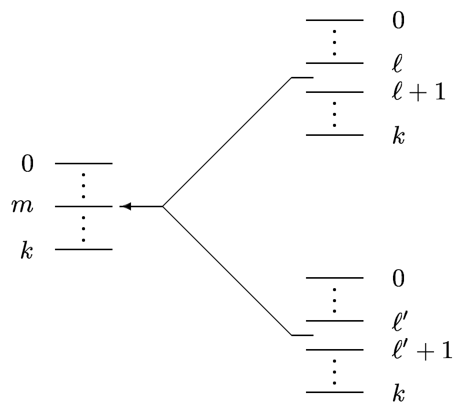
Approximation Algorithm for Asian Options (continued)

- But $A_{\rm u}$ is not likely to be one of the $k+1$ running averages at $N(j+1,i)$!
- Find the running averages that bracket it, that is,

$$A_\ell(j+1,i) \leq A_{\rm u} \leq A_{\ell+1}(j+1,i).$$

- Express $A_{\rm u}$ as a linearly interpolated value of the two running averages,

$$A_{\rm u} = x A_\ell(j+1,i) + (1-x)\, A_{\ell+1}(j+1,i), \quad 0 \leq x \leq 1.$$



Approximation Algorithm for Asian Options (continued)

- Obtain the approximate option value given the running average A_u via

$$C_u \equiv x C_\ell(j+1, i) + (1-x)\, C_{\ell+1}(j+1, i).$$

  - $C_\ell(t,s)$ denotes the option value at node $N(t,s)$ with running average $A_\ell(t,s)$.
- This interpolation introduces the second source of error.

Approximation Algorithm for Asian Options (concluded)

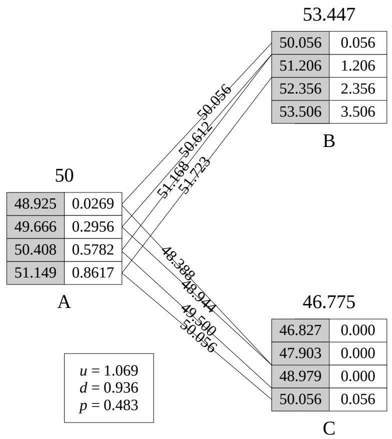
- The same steps are repeated for the down node $N(j+1, i+1)$ to obtain another approximate option value C_d.
- Finally obtain the option value as

$$(pC_u + (1-p)\, C_d)\, e^{-r\Delta t}.$$

- The running time is $O(kn^2)$.
  - There are $O(n^2)$ nodes.
  - Each node has $O(k)$ buckets.

A Numerical Example

- Consider a European arithmetic average-rate call with strike price 50.
- Assume zero interest rate in order to dispense with discounting.
- The minimum running average at node A in the figure on p. 149 is 48.925.
- The maximum running average at node A in the same figure is 51.149.



A Numerical Example (continued)

- Each node picks $k = 3$ for 4 equally spaced running averages.
- The same calculations are done for node A's successor nodes B and C.
- Suppose node A is 2 periods from the root node.
- Consider the up move from node A with running average 49.666.

A Numerical Example (continued)

- Because the stock price at node B is 53.447, the new running average will be

$$\frac{3 \times 49.666 + 53.447}{4} \approx 50.612.$$

- With 50.612 lying between 50.056 and 51.206 at node B, we solve

$$50.612 = x \times 50.056 + (1 - x) \times 51.206$$

to obtain $x \approx 0.517$.

A Numerical Example (continued)

- The option value corresponding to running average 51.206 at node B is 1.206.
- The option value corresponding to running average 50.056 at node B is 0.056.
- Their contribution to the option value corresponding to running average 49.666 at node A is weighted linearly as

$$x \times 0.056 + (1 - x) \times 1.206 \approx 0.611.$$

A Numerical Example (continued)

- Now consider the down move from node A with running average 49.666.
- Because the stock price at node C is 46.775, the new running average will be

$$\frac{3 \times 49.666 + 46.775}{4} \approx 48.944.$$

- With 48.944 lying between 47.903 and 48.979 at node C, we solve

$$48.944 = x \times 47.903 + (1 - x) \times 48.979$$

to obtain $x \approx 0.033$.

A Numerical Example (concluded)

- The option values corresponding to running averages 47.903 and 48.979 at node C are both 0.0.
- Their contribution to the option value corresponding to running average 49.666 at node A is 0.0.
- Finally, the option value corresponding to running average 49.666 at node A equals

$$p \times 0.611 + (1 - p) \times 0.0 \approx 0.2956,$$

where $p = 0.483$.

- The remaining three option values at node A can be computed similarly.

Hedging

Delta Hedge

- The delta (hedge ratio) of a derivative f is defined as $\Delta \equiv \partial f / \partial S$.
- Thus $\Delta f \approx \Delta \times \Delta S$ for relatively small changes in the stock price, ΔS.
- A delta-neutral portfolio is hedged as it is immunized against small changes in the stock price.

Delta Hedge (concluded)

- A trading strategy that dynamically maintains a delta-neutral portfolio is called delta hedge.
- Delta changes with the stock price.
- A delta hedge needs to be rebalanced periodically in order to maintain delta neutrality.

Implementing Delta Hedge

- We want to hedge N short derivatives.
- Assume the stock pays no dividends.
- The delta-neutral portfolio maintains $N \times \Delta$ shares of stock plus B borrowed dollars such that

$$-N \times f + N \times \Delta \times S - B = 0.$$

Implementing Delta Hedge (concluded)

- At next rebalancing point when the delta is Δ', buy $N \times (\Delta' - \Delta)$ shares to maintain $N \times \Delta'$ shares with a total borrowing of $B' = N \times \Delta' \times S' - N \times f'$.
- Delta hedge is the discrete-time analog of the continuous-time limit.
- It will rarely be self-financing.

The Setup

- A hedger is short 10,000 European calls.
- $\sigma = 30\%$; $r = 6\%$.
- This call's expiration is four weeks away, its strike price is \$50, and each call has a current value of $f = 1.76791$.
- As an option covers 100 shares of stock, $N = 1{,}000{,}000$.
- The trader adjusts the portfolio weekly.
- The calls are replicated well if the cumulative cost of trading stock is close to the call premium's FV.

Example

- As $\Delta = 0.538560$, $N \times \Delta = 538,560$ shares are purchased for a total cost of $538{,}560 \times 50 = 26{,}928{,}000$ dollars to make the portfolio delta-neutral.
- The trader finances the purchase by borrowing

$$B = N \times \Delta \times S - N \times f = 25{,}160{,}090$$

  dollars net.
- The portfolio has zero net value now.

Example (continued)

- At 3 weeks to expiration, the stock price rises to \$51.
- The new call value is $f' = 2.10580$.
- So the portfolio is worth

$$-N \times f' + 538{,}560 \times 51 - Be^{0.06/52} = 171,622$$

  before rebalancing.
- Delta hedge does not replicate the calls perfectly; it is not self-financing as \$171,622 can be withdrawn.

Example (continued)

- The magnitude of the tracking error—the variation in the net portfolio value—can be mitigated if adjustments are made more frequently.
- In fact, the tracking error is positive about 68% of the time even though its expected value is essentially zero.[a]
- It is furthermore proportional to vega.

[a] Boyle and Emanuel (1980).

Example (continued)

- In practice tracking errors will cease to decrease beyond a certain rebalancing frequency.
- With a higher delta $\Delta' = 0.640355$, the trader buys $N \times (\Delta' - \Delta) = 101,795$ shares for \$5,191,545.
- The number of shares is increased to $N \times \Delta' = 640,355$.

Example (continued)

- The cumulative cost is

$$26{,}928{,}000 \times e^{0.06/52} + 5{,}191{,}545 = 32{,}150{,}634.$$

- The net borrowed amount is

$$B' = 640{,}355 \times 51 - N \times f' = 30{,}552{,}305.$$

  - Alternatively, the number could be arrived at via

$$Be^{0.06/52} + 5{,}191{,}545 + 171{,}622 = 30{,}552{,}305.$$

- The portfolio is again delta-neutral with zero value.

τ	S (1)	Option value f (2)	Delta Δ (3)	Change in delta (5)	No. shares bought $N\times(5)$ (6)	Cost of shares $(1)\times(6)$ (7)	Cumulative cost FV(8')+(7) (8)
4	50	1.7679	0.53856	—	538,560	26,928,000	26,928,000
3	51	2.1058	0.64036	0.10180	101,795	5,191,545	32,150,634
2	53	3.3509	0.85578	0.21542	215,425	11,417,525	43,605,277
1	52	2.2427	0.83983	−0.01595	−15,955	−829,660	42,825,960
0	54	4.0000	1.00000	0.16017	160,175	8,649,450	51,524,853

The total number of shares is 1,000,000 at expiration (trading takes place at expiration, too).

Example (concluded)

- At expiration, the trader has 1,000,000 shares.
- They are exercised against by the in-the-money calls for $50,000,000.
- The trader is left with an obligation of

$$51{,}524{,}853 - 50{,}000{,}000 = 1{,}524{,}853,$$

  which represents the replication cost.
- Compared with the FV of the call premium,

$$1{,}767{,}910 \times e^{0.06 \times 4/52} = 1{,}776{,}088,$$

  the net gain is $1{,}776{,}088 - 1{,}524{,}853 = 251{,}235$.

Trees

I love a tree more than a man.
— Ludwig van Beethoven (1770–1827)

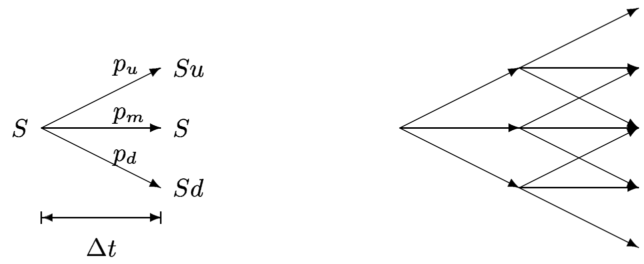
Trinomial Tree

- The three stock prices at time Δt are S, Su, and Sd, where $ud = 1$ (see next page).
- Let

$$\begin{aligned} M &\equiv e^{r\Delta t}, \\ V &\equiv M^2(e^{\sigma^2\Delta t} - 1). \end{aligned}$$

- Then

$$\begin{aligned} p_u &= \frac{u\left(V + M^2 - M\right) - (M-1)}{(u-1)\,(u^2-1)}, \\ p_d &= \frac{u^2\left(V + M^2 - M\right) - u^3(M-1)}{(u-1)\,(u^2-1)}. \end{aligned}$$



Trinomial Tree (concluded)

- Use $u = e^{\lambda\sigma\sqrt{\Delta t}}$, where $\lambda \geq 1$ is a tunable parameter.
- Then

$$\begin{aligned} p_u &\to \frac{1}{2\lambda^2} + \frac{\left(r + \sigma^2\right)\sqrt{\Delta t}}{2\lambda\sigma}, \\ p_d &\to \frac{1}{2\lambda^2} - \frac{\left(r - 2\sigma^2\right)\sqrt{\Delta t}}{2\lambda\sigma}. \end{aligned}$$

- A nice choice for λ is $\sqrt{\pi/2}$.

Barrier Options Revisited

- BOPM introduces a specification error by replacing the barrier with a nonidentical effective barrier.
- The trinomial model solves the problem by adjusting λ so that the barrier is hit exactly.[a]
- It takes

$$h = \frac{\ln(S/H)}{\lambda\sigma\sqrt{\Delta t}}$$

consecutive down moves to go from S to H if h is an integer, which is easy to achieve by adjusting λ.

[a] Ritchken (1995).

Barrier Options Revisited (continued)

- Typically, we find the smallest $\lambda \geq 1$ such that h is an integer, that is,

$$\lambda = \max_{j=1,2,3,\ldots} \frac{\ln(S/H)}{j\sigma\sqrt{\Delta t}}.$$

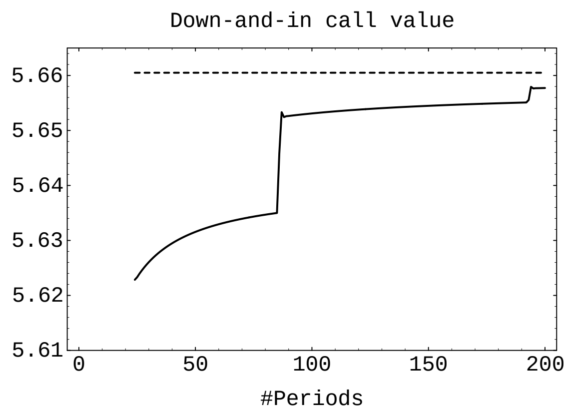
  - Such a λ may not exist for very small n's.
- This done, one of the layers of the trinomial tree coincides with the barrier.

Barrier Options Revisited (concluded)

- The following probabilities may be used,

$$\begin{aligned} p_u &= \frac{1}{2\lambda^2} + \frac{\mu'\sqrt{\Delta t}}{2\lambda\sigma}, \\ p_m &= 1 - \frac{1}{\lambda^2}, \\ p_d &= \frac{1}{2\lambda^2} - \frac{\mu'\sqrt{\Delta t}}{2\lambda\sigma}. \end{aligned}$$

  - $\mu' \equiv r - \sigma^2/2$.



Time Series Analysis

Conditional Variance Models for Price Volatility

- For many models, past information thus has no effect on the variance of prediction.
  - The Black-Scholes model is one example.
- To address this drawback, consider models for returns consistent with a changing conditional variance have been proposed.

ARCH[a] and GARCH Models

- An influential model in this direction is the autoregressive conditional heteroskedastic (ARCH) model.
- A very popular extension of the ARCH model is the generalized autoregressive conditional heteroskedastic (GARCH) process.

[a]Engle (1982), co-winner of the 2003 Nobel Prize in Economic Sciences.

ARCH and GARCH Models (continued)

- The simplest GARCH$(1,1)$ process is

$$V_t^2 = a_0 + a_1 (X_{t-1} - \mu)^2 + a_2 V_{t-1}^2.$$

- The volatility at time t as estimated at time $t-1$ depends on the squared return and the estimated volatility at time $t-1$.
- The estimate of volatility averages past squared returns by giving heavier weights to recent squared returns.

ARCH and GARCH Models (concluded)

- It is usually assumed that $a_1 + a_2 < 1$ and $a_0 > 0$, in which case the unconditional, long-run variance is given by $a_0/(1 - a_1 - a_2)$.
- A popular special case of GARCH(1, 1) is the exponentially weighted moving average process, which sets a_0 to zero and a_2 to $1 - a_1$.
- This model is used in J.P. Morgan's RiskMetrics™.

GARCH Option Pricing

- Options can be priced when the underlying asset's return follows a GARCH process.
- Let S_t denote the asset price at date t.
- Let h_t^2 be the conditional variance of the return over the period $[\,t, t+1\,]$ given the information at date t.
  - "One day" is merely a convenient term for any elapsed time Δt.

GARCH Option Pricing (continued)

- Adopt the following risk-neutral process for the price dynamics (Duan, 1995):

$$\ln \frac{S_{t+1}}{S_t} = r - \frac{h_t^2}{2} + h_t \epsilon_{t+1}, \tag{15}$$

where

$$\begin{aligned} h_{t+1}^2 &= \beta_0 + \beta_1 h_t^2 + \beta_2 h_t^2 (\epsilon_{t+1} - c)^2, \qquad (16) \\ \epsilon_{t+1} &\sim N(0,1) \text{ given information at date } t, \\ r &= \text{daily riskless return}, \\ c &\geq 0. \end{aligned}$$

GARCH Option Pricing (concluded)

- With $y_t \equiv \ln S_t$ denoting the logarithmic price, the model becomes

$$y_{t+1} = y_t + r - \frac{h_t^2}{2} + h_t \epsilon_{t+1}. \tag{17}$$

- The pair (y_t, h_t^2) completely describes the current state.
- The conditional mean and variance of y_{t+1} are

$$\begin{aligned} E[\, y_{t+1} \,|\, y_t, h_t^2 \,] &= y_t + r - \frac{h_t^2}{2}, \qquad (18) \\ \text{Var}[\, y_{t+1} \,|\, y_t, h_t^2 \,] &= h_t^2. \qquad (19) \end{aligned}$$

The Ritchken-Trevor (RT) Algorithm[a]

- The GARCH model is a continuous-state model.
- To approximate it, we turn to trees with *discrete* states.
- Path dependence in GARCH makes the tree for asset prices explode exponentially.
- We need to mitigate this combinatorial explosion somewhat.

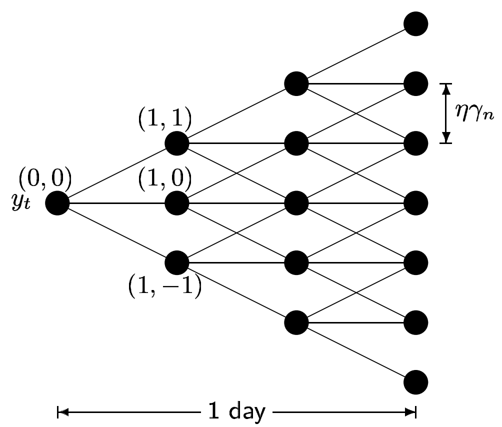
[a] Ritchken and Trevor (1999).

The Ritchken-Trevor Algorithm (continued)

- Partition a day into n periods.
- Three states follow each state (y_t, h_t^2) after a period.
- As the trinomial model combines, $2n+1$ states at date $t+1$ follow each state at date t (recall p. 171).
- These $2n+1$ values must approximate the probability distribution of (y_{t+1}, h_{t+1}^2).
- So the conditional moments (18)–(19) at date $t+1$ on p. 184 must be matched by the trinomial model to guarantee convergence to the continuous-state model.

The Ritchken-Trevor Algorithm (continued)

- It remains to pick the jump size and the three branching probabilities.
- Define $\gamma \equiv h_0$ (other multiples of h_0 are possible).
- Define
$$\gamma_n \equiv \frac{\gamma}{\sqrt{n}}.$$
- The jump size will be some integer multiple η of γ_n.
- We call η the jump parameter (see p. 188).



The seven values on the right approximate the distribution of logarithmic price y_{t+1}.

The Ritchken-Trevor Algorithm (continued)

- The middle branch does not change the underlying asset's price.
- The probabilities for the up, middle, and down branches are

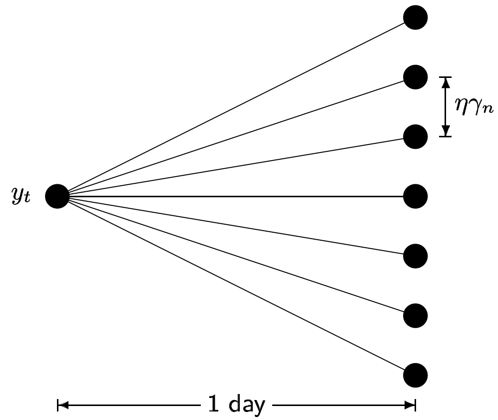
$$p_u = \frac{h_t^2}{2\eta^2\gamma^2} + \frac{r - (h_t^2/2)}{2\eta\gamma\sqrt{n}}, \quad (20)$$

$$p_m = 1 - \frac{h_t^2}{\eta^2\gamma^2}, \quad (21)$$

$$p_d = \frac{h_t^2}{2\eta^2\gamma^2} - \frac{r - (h_t^2/2)}{2\eta\gamma\sqrt{n}}. \quad (22)$$

The Ritchken-Trevor Algorithm (continued)

- We can dispense with the intermediate nodes *between* dates to create a $(2n+1)$-nomial tree (see p. 191).
- The resulting model is multinomial with $2n+1$ branches from any state (y_t, h_t^2).
- There are two reasons behind this manipulation.
  - Interdate nodes are created merely to approximate the continuous-state model after one day.
  - Keeping the interdate nodes results in a tree that is n times as large.



This heptanomial tree is the outcome of the trinomial tree on p. 188 after its intermediate nodes are removed.

The Ritchken-Trevor Algorithm (continued)

- A node with logarithmic price $y_t + \ell\eta\gamma_n$ at date $t+1$ follows the current node at date t with price y_t for some $-n \leq \ell \leq n$.
- The probability that this happens is

$$P(\ell) \equiv \sum_{j_u, j_m, j_d} \frac{n!}{j_u!\, j_m!\, j_d!}\, p_u^{j_u} p_m^{j_m} p_d^{j_d},$$

  with $j_u, j_m, j_d \geq 0$, $n = j_u + j_m + j_d$, and $\ell = j_u - j_d$.
- They can be computed in $O(n^2)$ time.

The Ritchken-Trevor Algorithm (continued)

- The updating rule (16) on p. 183 must be modified to account for the adoption of the discrete-state model.
- The logarithmic price $y_t + \ell \eta \gamma_n$ at date $t+1$ following state (y_t, h_t^2) at date t has a variance equal to

$$h_{t+1}^2 = \beta_0 + \beta_1 h_t^2 + \beta_2 h_t^2 (\epsilon'_{t+1} - c)^2, \qquad (23)$$

  - Above,

$$\epsilon'_{t+1} = \frac{\ell \eta \gamma_n - (r - h_t^2/2)}{h_t}, \quad \ell = 0, \pm 1, \pm 2, \ldots, \pm n,$$

    is a discrete random variable with $2n+1$ values.

The Ritchken-Trevor Algorithm (continued)

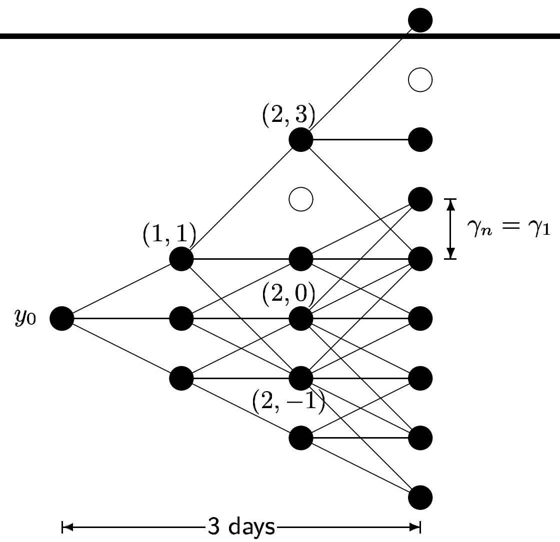
- Different conditional variances h_t^2 may require different η so that the probabilities calculated by Eqs. (20)–(22) on p. 189 lie between 0 and 1.
- This implies varying jump sizes.
- The necessary requirement $p_m \geq 0$ implies $\eta \geq h_t/\gamma$.
- Hence we try

$$\eta = \lceil h_t/\gamma \rceil, \lceil h_t/\gamma \rceil + 1, \lceil h_t/\gamma \rceil + 2, \ldots$$

  until valid probabilities are obtained or until their nonexistence is confirmed.

The Ritchken-Trevor Algorithm (continued)

- Obviously, the magnitude of η tends to grow with h_t.
- The plot on p. 196 uses $n = 1$ to illustrate our points for a 3-day model.
- For example, node $(1, 1)$ of date 1 and node $(2, 3)$ of date 2 pick $\eta = 2$.



The Ritchken-Trevor Algorithm (continued)

- The topology of the tree is not a standard combining multinomial tree.
- For example, a few nodes on p. 196 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^2_{max}) and (y_t, h^2_{min}).
- Each of (y_t, h^2_{max}) and (y_t, h^2_{min}) 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]

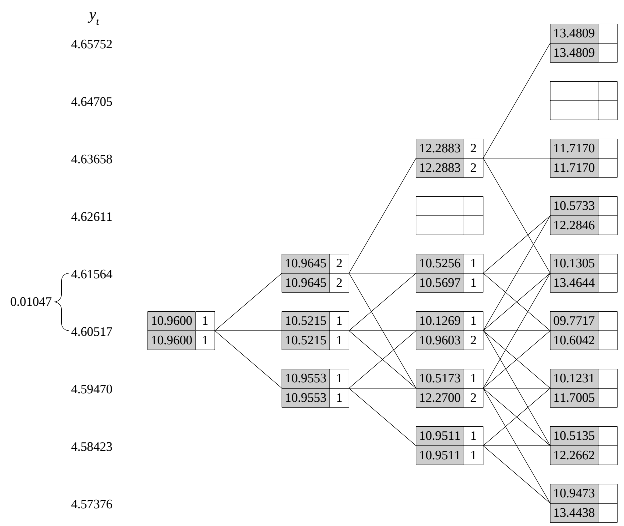
[a] Lyuu and Wu (2003).
[b] Lyuu and Wu (2003).

Numerical Examples

- $S_0 = 100$; $y_0 = \ln S_0 = 4.60517$; $r = 0$; $h_0^2 = 0.0001096$; $\gamma = h_0 = 0.010469$; $n = 1$; $\gamma_n = \gamma/\sqrt{n} = 0.010469$; $\beta_0 = 0.000006575$; $\beta_1 = 0.9$; $\beta_2 = 0.04$; $c = 0$.
  - 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 3-day tree is depicted on p. 202.



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. (20)–(22) on p. 189 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 (23) on p. 193 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. (20)–(22) on p. 189 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 (23) on p. 193.
- Carry out similar calculations for node $(1,-1)$ with $\ell = -1$ in updating rule (23).
- 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 (23) on p. 193 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 (23) on p. 193 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. (20)–(22) on p. 189 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. (20)–(22) on p. 189 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 (23) on p. 193 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 (23) on p. 193 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 (23) on p. 193 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. (20)–(22) on p. 189 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. (20)–(22) on p. 189 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.

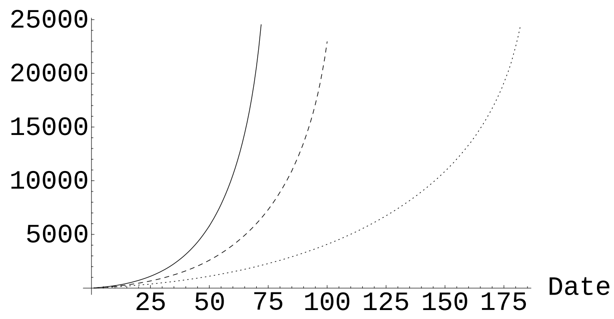
Negative Aspects of the RT Algorithm Revisited[a]

- Recall the problems mentioned on p. 199.
- 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. 202.
- In general, the kth variance at node (i, j) is

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

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

- 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. 142, where we dealt with arithmetic average-rate options.

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

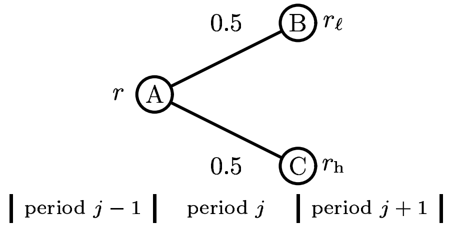
- A stochastic interest rate model performs two tasks.
  - Provides a stochastic process that defines future term structures without arbitrage profits.
  - "Consistent" with the observed term structures.

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. 229 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_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_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

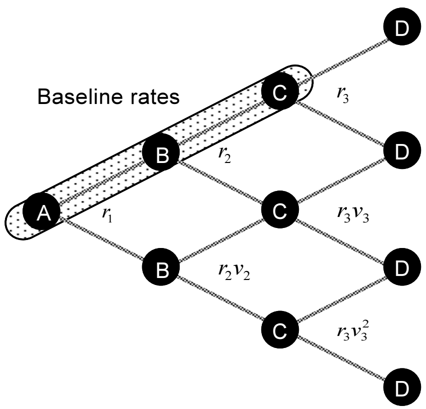
$$\sigma = \frac{1}{2} \frac{1}{\sqrt{\Delta t}} \ln \left(\frac{r_\mathrm{h}}{r_\ell} \right).$$

Binomial Interest Rate Tree (continued)

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

Binomial Interest Rate Tree (concluded)

- 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
$$v_j \equiv e^{2\sigma_j \sqrt{\Delta t}} \tag{24}$$
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.

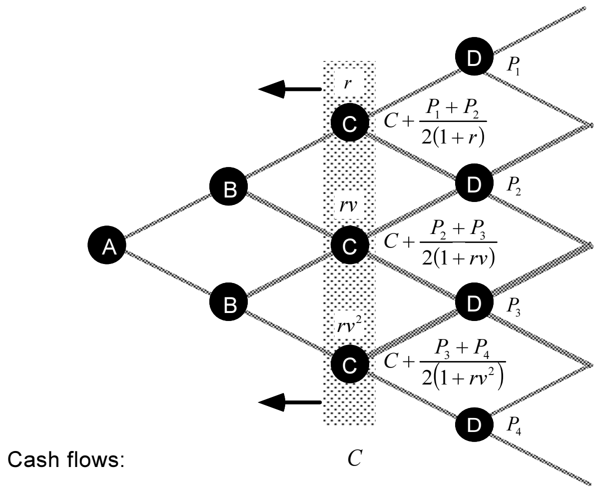


Set Things in Motion

- 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., the on-the-run yield curve obtained by the most recently issued Treasury securities may be used as the benchmark curve.
- The binomial tree should be consistent with both term structures.
- Here we focus on the term structure of interest rates.

Model Term Structures

- The model price is computed by backward induction.
- Refer back to the figure on p. 229.
- 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
$$\frac{P_{\rm B} + P_{\rm C}}{2(1+r)} + \text{cash flow at node A.}$$
- We compute the values column by column without explicitly expanding the binomial interest rate tree.
- This takes quadratic time and linear space.



Sample Term Structure

- We shall construct interest rate trees consistent with the sample term structure in the following table.
- Assume the short rate volatility is such that $v \equiv r_{\mathrm{h}}/r_{\ell} = 1.5$, independent of time.

Period	1	2	3
Spot rate (%)	4	4.2	4.3
One-period forward rate (%)	4	4.4	4.5
Discount factor	0.96154	0.92101	0.88135

Issues in Calibration

- The model prices generated by the binomial interest rate tree should match the observed market prices.
- Treat the backward induction for the model price of the m-period zero-coupon bond as computing some function of the unknown baseline rate r_m called $f(r_m)$.
- A root-finding method is applied to solve $f(r_m) = P$ for r_m given the zero's price P and $r_1, r_2, \ldots, r_{m-1}$.
- This procedure is carried out for $m = 1, 2, \ldots, n$.
- Runs in cubic time, hopelessly slow.

Binomial Interest Rate Tree Calibration

- Calibration can be accomplished in quadratic time by the use of forward induction (Jamshidian, 1991).
- The scheme records how much \$1 at a node contributes to the model price.
- This number is called the state price.
  - It stands for the price of a claim that pays \$1 at that particular node (state) and 0 elsewhere.
- The column of state prices will be established by moving *forward* from time 1 to time n.

Binomial Interest Rate Tree Calibration (continued)

- Suppose we are at time j and there are $j+1$ nodes.
  - The baseline rate for period j is $r \equiv r_j$.
  - The multiplicative ratio be $v \equiv v_j$.
  - $P_1, P_2, \ldots, P_j$ are the state prices a period prior, corresponding to rates $r, rv, \ldots, rv^{j-1}$.
- By definition, $\sum_{i=1}^{j} P_i$ is the price of the $(j-1)$-period zero-coupon bond.

Binomial Interest Rate Tree Calibration (continued)

- One dollar at time j has a known market value of $1/[\,1+S(j)\,]^j$, where $S(j)$ is the j-period spot rate.
- Alternatively, this dollar has a present value of

$$g(r) \equiv \frac{P_1}{(1+r)} + \frac{P_2}{(1+rv)} + \frac{P_3}{(1+rv^2)} + \cdots + \frac{P_j}{(1+rv^{j-1})}.$$

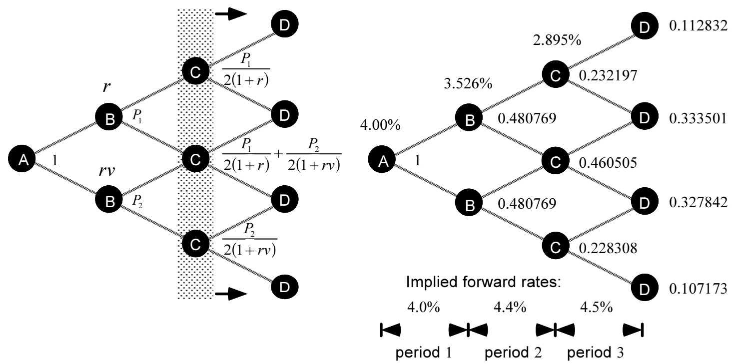
- So we solve

$$g(r) = \frac{1}{[\,1+S(j)\,]^j} \tag{25}$$

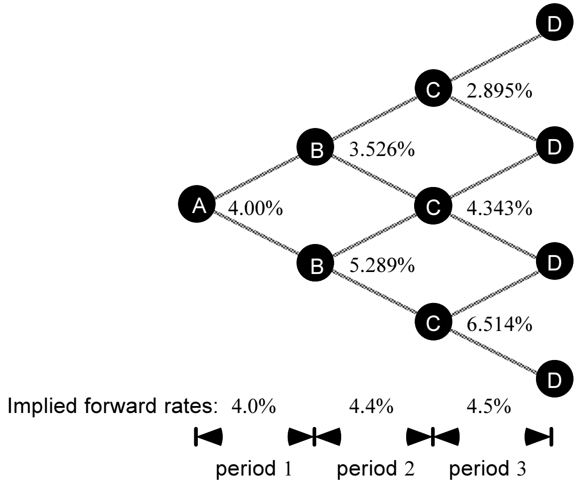
for r.

Binomial Interest Rate Tree Calibration (continued)

- Given a decreasing market discount function, a unique positive solution for r is guaranteed.
- The state prices at time j can now be calculated (see figure (a) next page).
- We call a tree with these state prices a binomial state price tree (see figure (b) next page).
- The calibrated tree is depicted in on p. 245.



(a) (b)



Binomial Interest Rate Tree Calibration (concluded)

- The Newton-Raphson method can be used to solve for the r in Eq. (25) on p. 242 as $g'(r)$ is easy to evaluate.
- The above idea is straightforward to implement.
- The total running time is $O(n^2)$.
- With a good initial guess, the Newton-Raphson method converges in only a few steps.[a]

[a] Lyuu (1999).

A Numerical Example

- One dollar at the end of the second period should have a present value of 0.92101 by the sample term structure.
- The baseline rate for the second period, r_2, satisfies
$$\frac{0.480769}{1+r_2}+\frac{0.480769}{1+1.5\times r_2}=0.92101.$$
- The result is $r_2 = 3.526\%$.
- This is used to derive the next column of state prices shown in figure (b) on p. 244 as 0.232197, 0.460505, and 0.228308.
- Their sum gives the market discount factor 0.92101.

A Numerical Example (concluded)

- The baseline rate for the third period, r_3, satisfies
$$\frac{0.232197}{1+r_3}+\frac{0.460505}{1+1.5\times r_3}+\frac{0.228308}{1+(1.5)^2\times r_3}=0.88135.$$
- The result is $r_3 = 2.895\%$.
- Now,
$$\frac{1}{4}\times\frac{1}{1.04}\times[\frac{1}{1.03526}\times(\frac{1}{1.02895}+\frac{1}{1.04343})+\frac{1}{1.05289}\times(\frac{1}{1.04343}+\frac{1}{1.06514})],$$
which equals 0.88135, an exact match.
- The tree on p. 245 prices without bias the benchmark securities.

Spread of Nonbenchmark Bonds

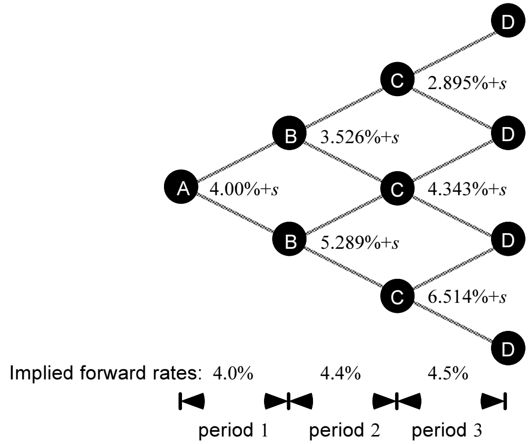
- Model prices calculated by the calibrated tree as a rule do not match market prices of nonbenchmark bonds.
- The incremental return over the benchmark bonds is called spread.
- We look for the spread that, when added *uniformly* over the short rates in the tree, makes the model price equal the market price.

Spread of Nonbenchmark Bonds (continued)

- We illustrate the idea with an example.
- Start with the tree on p. 251.
- Consider a security with cash flow C_i at time i for $i = 1, 2, 3$.
- Its model price is $p(s)$, which is equal to

$$\frac{1}{1.04+s} \times \left[C_1 + \frac{1}{2} \times \frac{1}{1.03526+s} \times \left(C_2 + \frac{1}{2} \left(\frac{C_3}{1.02895+s} + \frac{C_3}{1.04343+s} \right) \right) + \frac{1}{2} \times \frac{1}{1.05289+s} \times \left(C_2 + \frac{1}{2} \left(\frac{C_3}{1.04343+s} + \frac{C_3}{1.06514+s} \right) \right) \right].$$

- Given a market price of P, the spread is the s that solves $P = p(s)$.



Spread of Nonbenchmark Bonds (continued)

- We will employ the Newton-Raphson root-finding method to solve $p(s) - P = 0$ for s.
- But a quick look at the equation above reveals that evaluating $p'(s)$ directly is infeasible.
- Fortunately, the tree can be used to evaluate both $p(s)$ and $p'(s)$ during backward induction.

Spread of Nonbenchmark Bonds (continued)

- Consider an arbitrary node A in the tree associated with the short rate r.
- In the process of computing the model price $p(s)$, a price $p_{\mathrm{A}}(s)$ is computed at A.
- Prices computed at A's two successor nodes B and C are discounted by $r + s$ to obtain $p_{\mathrm{A}}(s)$ as follows,

$$p_{\mathrm{A}}(s) = c + \frac{p_{\mathrm{B}}(s) + p_{\mathrm{C}}(s)}{2(1 + r + s)},$$

where c denotes the cash flow at A.

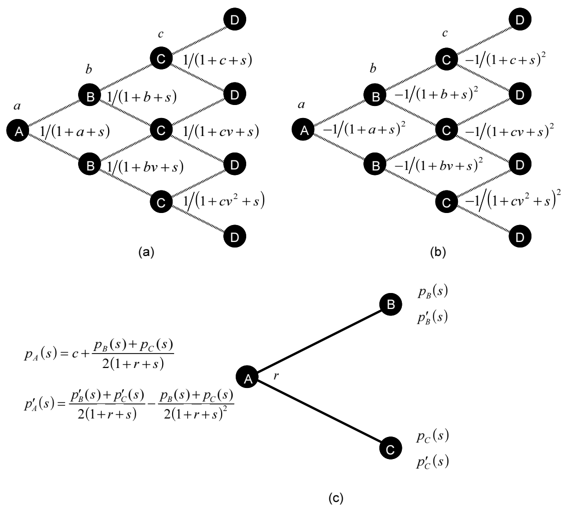
Spread of Nonbenchmark Bonds (continued)

- To compute $p'_{\mathrm{A}}(s)$ as well, node A calculates

$$p'_{\mathrm{A}}(s) = \frac{p'_{\mathrm{B}}(s) + p'_{\mathrm{C}}(s)}{2(1 + r + s)} - \frac{p_{\mathrm{B}}(s) + p_{\mathrm{C}}(s)}{2(1 + r + s)^2}. \quad (26)$$

- This is easy if $p'_{\mathrm{B}}(s)$ and $p'_{\mathrm{C}}(s)$ are also computed at nodes B and C.
- Apply the above procedure inductively to yield $p(s)$ and $p'(s)$ at the root (see p. 255).
- This is called the differential tree method.[a]

[a] Lyuu (1999).



Spread of Nonbenchmark Bonds (continued)

- The total running time is $O(n^2)$.
- The memory requirement is $O(n)$.

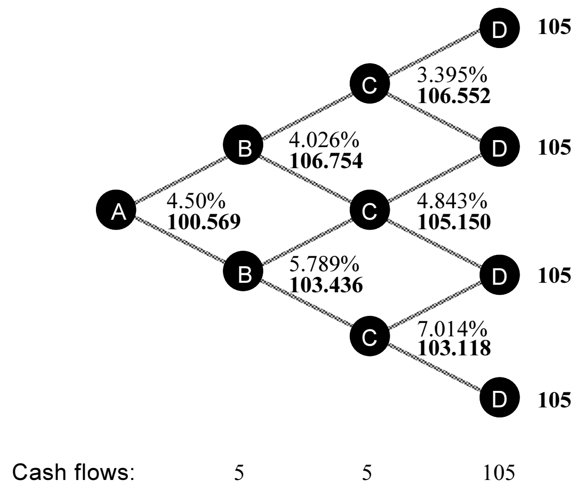
Spread of Nonbenchmark Bonds (continued)

Number of partitions n	Running time (s)	Number of iterations	Number of partitions	Running time (s)	Number of iterations
500	7.850	5	10500	3503.410	5
1500	71.650	5	11500	4169.570	5
2500	198.770	5	12500	4912.680	5
3500	387.460	5	13500	5714.440	5
4500	641.400	5	14500	6589.360	5
5500	951.800	5	15500	7548.760	5
6500	1327.900	5	16500	8502.950	5
7500	1761.110	5	17500	9523.900	5
8500	2269.750	5	18500	10617.370	5
9500	2834.170	5	....	....	....

75MHz Sun SPARCstation 20.

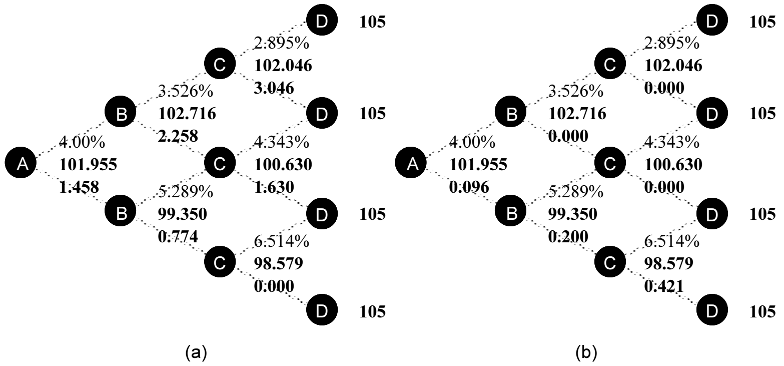
Spread of Nonbenchmark Bonds (concluded)

- Consider a three-year, 5% bond with a market price of 100.569.
- Assume the bond pays annual interest.
- The spread can be shown to be 50 basis points over the tree (see p. 259).
- Note that the idea of spread does not assume parallel shifts in the term structure.



Fixed-Income Options

- Consider a two-year 99 European call on the three-year, 5% Treasury.
- Assume the Treasury pays annual interest.
- From p. 261 the three-year Treasury's price minus the $5 interest could be $102.046, $100.630, or $98.579 two years from now.
- Since these prices do not include the accrued interest, we should compare the strike price against them.
- The call is therefore in the money in the first two scenarios, with values of $3.046 and $1.630, and out of the money in the third scenario.



Fixed-Income Options (concluded)

- The option value is calculated to be \$1.458 on p. 261(a).
- European interest rate puts can be valued similarly.
- Consider a two-year 99 European put on the same security.
- At expiration, the put is in the money only if the Treasury is worth \$98.579 without the accrued interest.
- The option value is computed to be \$0.096 on p. 261(b).

Finis