Time-Varying Double Barriers under Time-Dependent Volatility

- More general models allow a time-varying $\sigma(t)$ (p. 312).
- Let the two barriers $L(t)$ and $H(t)$ be functions of time.$^b$
- They do not have to be differentiable or even continuous.
- Still, we can price double-barrier options in $O(n^2)$ time or less with trinomial trees.
- Continuously monitored double-barrier knock-out options with time-varying barriers are called hot dog options.$^c$

---

$^a$Y. Zhang (R05922052) (2019).
$^b$So the barriers are continuously monitored.
General Local-Volatility Models and Their Trees

- Consider the general local-volatility model
  \[
  \frac{dS}{S} = (r_t - q_t) \, dt + \sigma(S, t) \, dW,
  \]
  where \( L \leq \sigma(S, t) \leq U \) for some positive \( L \) and \( U \).

- This model has a unique (weak) solution.\(^a\)

- The positive lower bound is justifiable because prices fluctuate.

\(^a\) Achdou & Pironneau (2005).
The upper-bound assumption is also reasonable.

Even on October 19, 1987, the CBOE S&P 100 Volatility Index (VXO) was about 150%, the highest ever.\(^a\)

An efficient quadratic-sized tree for this range-bounded model is straightforward.\(^b\)

Pick any \(\sigma' \geq U\).

Grow the trinomial tree with the node spacing \(\sigma' \sqrt{\Delta t}\).\(^c\)

The branching probabilities are guaranteed to be valid.

\(^a\)Caprio (2012).
\(^b\)Lok (D99922028) & Lyuu (2016, 2017, 2020).
\(^c\)Haahtela (2010).
General Local-Volatility Models and Their Trees (concluded)

- The same idea can be applied to price double-barrier options.
- Pick any
  \[ \sigma' \geq \max \left[ \max_{S_0 \leq t \leq T} \sigma(S, t), \sqrt{2} \sigma(S_0, 0) \right]. \]
- Grow the trinomial tree with the node spacing \( \sigma' \sqrt{\Delta t} \).
- Apply the mean-tracking idea to the first period and Eqs. (100)–(105) on p. 760 to obtain the probabilities
Merton’s Jump-Diffusion Model

• Empirically, stock returns tend to have fat tails, inconsistent with the Black-Scholes model’s assumptions.

• Stochastic volatility and jump processes have been proposed to address this problem.

• Merton’s (1976) jump-diffusion model is our focus.
Merton’s Jump-Diffusion Model (continued)

• This model superimposes a jump component on a diffusion component.

• The diffusion component is the familiar geometric Brownian motion.

• The jump component is composed of lognormal jumps driven by a Poisson process.
  – It models the rare but large changes in the stock price because of the arrival of important new information.\(^a\)

---
\(^a\)Derman & M. B. Miller (2016), “There is no precise, universally accepted definition of a jump, but it usually comes down to magnitude, duration, and frequency.”
Merton’s Jump-Diffusion Model (continued)

• Let $S_t$ be the stock price at time $t$.

• The risk-neutral jump-diffusion process for the stock price follows\(^\text{a}\)

\[
\frac{dS_t}{S_t} = (r - \lambda \bar{k}) dt + \sigma dW_t + k dq_t. \tag{107}
\]

• Above, $\sigma$ denotes the volatility of the diffusion component.

\(^{a}\)Derman & M. B. Miller (2016), “[M]ost jump-diffusion models simply assume risk-neutral pricing without convincing justification.”
Merton’s Jump-Diffusion Model (continued)

• The jump event is governed by a compound Poisson process \( q_t \) with intensity \( \lambda \), where \( k \) denotes the magnitude of the random jump.
  
  – The distribution of \( k \) obeys

\[
\ln(1 + k) \sim N(\gamma, \delta^2)
\]

with mean \( \bar{k} \triangleq E(k) = e^{\gamma + \delta^2/2} - 1 \).

– Note that \( k > -1 \).

– Note also that \( k \) is not related to \( dt \).

• The model with \( \lambda = 0 \) reduces to the Black-Scholes model.
Merton’s Jump-Diffusion Model (continued)

- The solution to Eq. (107) on p. 798 is

\[ S_t = S_0 e^{(r-\lambda \bar{k}-\sigma^2/2)t} + \sigma W_t U(n(t)), \tag{108} \]

where

\[ U(n(t)) = \prod_{i=0}^{n(t)} (1 + k_i). \]

- \( k_i \) is the magnitude of the \( i \)th jump with \( \ln(1 + k_i) \sim N(\gamma, \delta^2) \).
- \( k_0 = 0 \).
- \( n(t) \) is a Poisson process with intensity \( \lambda \).
Merton’s Jump-Diffusion Model (concluded)

- Recall that \( n(t) \) denotes the number of jumps that occur up to time \( t \).
- It is known that \( E[n(t)] = \text{Var}[n(t)] = \lambda t \).
- As \( k_i > -1 \), stock prices will stay positive.
- The geometric Brownian motion, the lognormal jumps, and the Poisson process are assumed to be independent.
Tree for Merton’s Jump-Diffusion Model\textsuperscript{a}

- Define the $S$-logarithmic return of the stock price $S'$ as
  \[ \ln(S'/S). \]

- Define the logarithmic distance between stock prices $S'$ and $S$ as
  \[ | \ln(S') - \ln(S) | = | \ln(S'/S) |. \]

\textsuperscript{a}Dai (B82506025, R86526008, D8852600), C. Wang (F95922018), Lyuu, & Y. Liu (2010).
Tree for Merton’s Jump-Diffusion Model (continued)

- Take the logarithm of Eq. (108) on p. 800:

\[ M_t \triangleq \ln \left( \frac{S_t}{S_0} \right) = X_t + Y_t, \tag{109} \]

where

\[ X_t \triangleq \left( r - \lambda \bar{k} - \frac{\sigma^2}{2} \right) t + \sigma W_t, \tag{110} \]

\[ Y_t \triangleq \sum_{i=0}^{n(t)} \ln (1 + k_i). \tag{111} \]

- It decomposes the \( S_0 \)-logarithmic return of \( S_t \) into the diffusion component \( X_t \) and the jump component \( Y_t \).
Tree for Merton’s Jump-Diffusion Model (continued)

- Motivated by decomposition (109) on p. 803, the tree construction divides each period into a diffusion phase followed by a jump phase.

- In the diffusion phase, $X_t$ is approximated by the BOPM.

- So $X_t$ makes an up move to $X_t + \sigma \sqrt{\Delta t}$ with probability $p_u$ or a down move to $X_t - \sigma \sqrt{\Delta t}$ with probability $p_d$. 
Tree for Merton’s Jump-Diffusion Model (continued)

• According to BOPM,

\[ p_u = \frac{e^{\mu \Delta t} - d}{u - d}, \]
\[ p_d = 1 - p_u, \]

except that \( \mu = r - \lambda \bar{k} \) here.

• The diffusion component gives rise to diffusion nodes.

• They are spaced at \( 2\sigma \sqrt{\Delta t} \) apart such as the white nodes A, B, C, D, E, F, and G on p. 806.
White nodes are \textit{diffusion nodes}. Gray nodes are \textit{jump nodes}. In the diffusion phase, the solid black lines denote the binomial structure of BOPM; the dashed lines denote the trinomial structure. Only the double-circled nodes will remain after the construction. Note that a and b are diffusion nodes because no jump occurs in the jump phase.
Tree for Merton’s Jump-Diffusion Model (continued)

- In the jump phase, $Y_{t+\Delta t}$ is approximated by moves from each diffusion node to $2m$ jump nodes that match the first $2m$ moments of the lognormal jump.

- The $m$ jump nodes above the diffusion node are spaced at $h \overset{\Delta}{=} \sqrt{\gamma^2 + \delta^2}$ apart.

- Note that $h$ is independent of $\Delta t$. 
Tree for Merton’s Jump-Diffusion Model (concluded)

• The same holds for the $m$ jump nodes below the diffusion node.

• The gray nodes at time $\ell \Delta t$ on p. 806 are jump nodes.
  – We set $m = 1$ on p. 806.

• The size of the tree is $O(n^{2.5})$. 
Multivariate Contingent Claims

• They depend on two or more underlying assets.

• The basket call on \( m \) assets has the terminal payoff

\[
\max \left( \sum_{i=1}^{m} \alpha_i S_i(\tau) - X, 0 \right),
\]

where \( \alpha_i \) is the percentage of asset \( i \).

• Basket options are essentially options on a portfolio of stocks (or index options).\(^a\)

• Option on the best of two risky assets and cash has a terminal payoff of \( \max(S_1(\tau), S_2(\tau), X) \).

\(^a\)Except that membership and weights do not change for basket options (Bennett, 2014).
<table>
<thead>
<tr>
<th>Name</th>
<th>Payoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exchange option</td>
<td>$\max(S_1(\tau) - S_2(\tau), 0)$</td>
</tr>
<tr>
<td>Better-off option</td>
<td>$\max(S_1(\tau), \ldots, S_k(\tau), 0)$</td>
</tr>
<tr>
<td>Worst-off option</td>
<td>$\min(S_1(\tau), \ldots, S_k(\tau), 0)$</td>
</tr>
<tr>
<td>Binary maximum option</td>
<td>$I{ \max(S_1(\tau), \ldots, S_k(\tau)) &gt; X }$</td>
</tr>
<tr>
<td>Maximum option</td>
<td>$\max(\max(S_1(\tau), \ldots, S_k(\tau)) - X, 0)$</td>
</tr>
<tr>
<td>Minimum option</td>
<td>$\max(\min(S_1(\tau), \ldots, S_k(\tau)) - X, 0)$</td>
</tr>
<tr>
<td>Spread option</td>
<td>$\max(S_1(\tau) - S_2(\tau) - X, 0)$</td>
</tr>
<tr>
<td>Basket average option</td>
<td>$\max((S_1(\tau) + \cdots + S_k(\tau))/k - X, 0)$</td>
</tr>
<tr>
<td>Multi-strike option</td>
<td>$\max(S_1(\tau) - X_1, \ldots, S_k(\tau) - X_k, 0)$</td>
</tr>
<tr>
<td>Pyramid rainbow option</td>
<td>$\max(</td>
</tr>
<tr>
<td>Madonna option</td>
<td>$\max(\sqrt{(S_1(\tau) - X_1)^2 + \cdots + (S_k(\tau) - X_k)^2} - X, 0)$</td>
</tr>
</tbody>
</table>

\(^{a}\)Lyuu & Teng (R91723054) (2011).
Correlated Trinomial Model$^a$

- Two risky assets $S_1$ and $S_2$ follow

$$\frac{dS_i}{S_i} = r \, dt + \sigma_i \, dW_i$$

in a risk-neutral economy, $i = 1, 2$.

- Let

$$M_i \triangleq e^{r \Delta t},$$

$$V_i \triangleq M_i^2(e^{\sigma_i^2 \Delta t} - 1).$$

- $S_i \, M_i$ is the mean of $S_i$ at time $\Delta t$.
- $S_i^2 \, V_i$ the variance of $S_i$ at time $\Delta t$.

$^a$Boyle, Evnine, & Gibbs (1989).
Correlated Trinomial Model (continued)

- The value of $S_1S_2$ at time $\Delta t$ has a joint lognormal distribution with mean $S_1S_2M_1M_2e^{\rho\sigma_1\sigma_2\Delta t}$, where $\rho$ is the correlation between $dW_1$ and $dW_2$.

- Next match the 1st and 2nd moments of the approximating discrete distribution to those of the continuous counterpart.

- At time $\Delta t$ from now, there are 5 distinct outcomes.
Correlated Trinomial Model (continued)

- The five-point probability distribution of the asset prices is

<table>
<thead>
<tr>
<th>Probability</th>
<th>Asset 1</th>
<th>Asset 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_1$</td>
<td>$S_1u_1$</td>
<td>$S_2u_2$</td>
</tr>
<tr>
<td>$p_2$</td>
<td>$S_1u_1$</td>
<td>$S_2d_2$</td>
</tr>
<tr>
<td>$p_3$</td>
<td>$S_1d_1$</td>
<td>$S_2d_2$</td>
</tr>
<tr>
<td>$p_4$</td>
<td>$S_1d_1$</td>
<td>$S_2u_2$</td>
</tr>
<tr>
<td>$p_5$</td>
<td>$S_1$</td>
<td>$S_2$</td>
</tr>
</tbody>
</table>

- As usual, impose $u_id_i = 1$. 

©2021 Prof. Yuh-Dauh Lyuu, National Taiwan University
Correlated Trinomial Model (continued)

• The probabilities must sum to one, and the means must be matched:

\[ 1 = p_1 + p_2 + p_3 + p_4 + p_5, \]
\[ S_1 M_1 = (p_1 + p_2) S_1 u_1 + p_5 S_1 + (p_3 + p_4) S_1 d_1, \]
\[ S_2 M_2 = (p_1 + p_4) S_2 u_2 + p_5 S_2 + (p_2 + p_3) S_2 d_2. \]
Correlated Trinomial Model (concluded)

• Let $R \triangleq M_1 M_2 e^{\rho \sigma_1 \sigma_2 \Delta t}$.

• Match the variances and covariance:

\[
\begin{align*}
S_1^2 V_1 &= (p_1 + p_2) \left[ (S_1 u_1)^2 - (S_1 M_1)^2 \right] + p_5 \left[ S_1^2 - (S_1 M_1)^2 \right] \\
&\quad + (p_3 + p_4) \left[ (S_1 d_1)^2 - (S_1 M_1)^2 \right], \\
S_2^2 V_2 &= (p_1 + p_4) \left[ (S_2 u_2)^2 - (S_2 M_2)^2 \right] + p_5 \left[ S_2^2 - (S_2 M_2)^2 \right] \\
&\quad + (p_2 + p_3) \left[ (S_2 d_2)^2 - (S_2 M_2)^2 \right], \\
S_1 S_2 R &= (p_1 u_1 u_2 + p_2 u_1 d_2 + p_3 d_1 d_2 + p_4 d_1 u_2 + p_5) S_1 S_2.
\end{align*}
\]

• The solutions appear on p. 246 of the textbook.
Correlated Trinomial Model Simplified

- Let \( \mu_i' \triangleq r - \sigma_i^2/2 \) and \( u_i \triangleq e^{\lambda \sigma_i \sqrt{\Delta t}} \) for \( i = 1, 2 \).

- The following simpler scheme is good enough:

\[
\begin{align*}
    p_1 &= \frac{1}{4} \left[ \frac{1}{\lambda^2} + \frac{\sqrt{\Delta t}}{\lambda} \left( \frac{\mu_1'}{\sigma_1} + \frac{\mu_2'}{\sigma_2} \right) + \frac{\rho}{\lambda^2} \right], \\
    p_2 &= \frac{1}{4} \left[ \frac{1}{\lambda^2} + \frac{\sqrt{\Delta t}}{\lambda} \left( \frac{\mu_1'}{\sigma_1} - \frac{\mu_2'}{\sigma_2} \right) - \frac{\rho}{\lambda^2} \right], \\
    p_3 &= \frac{1}{4} \left[ \frac{1}{\lambda^2} + \frac{\sqrt{\Delta t}}{\lambda} \left( -\frac{\mu_1'}{\sigma_1} - \frac{\mu_2'}{\sigma_2} \right) + \frac{\rho}{\lambda^2} \right], \\
    p_4 &= \frac{1}{4} \left[ \frac{1}{\lambda^2} + \frac{\sqrt{\Delta t}}{\lambda} \left( -\frac{\mu_1'}{\sigma_1} + \frac{\mu_2'}{\sigma_2} \right) - \frac{\rho}{\lambda^2} \right], \\
    p_5 &= 1 - \frac{1}{\lambda^2}.
\end{align*}
\]

\(^a\text{Madan, Milne, & Shefrin (1989).}\)
Correlated Trinomial Model Simplified (continued)

• All of the probabilities lie between 0 and 1 if and only if
  
  \[ -1 + \lambda \sqrt{\Delta t} \left| \frac{\mu'_1}{\sigma_1} + \frac{\mu'_2}{\sigma_2} \right| \leq \rho \leq 1 - \lambda \sqrt{\Delta t} \left| \frac{\mu'_1}{\sigma_1} - \frac{\mu'_2}{\sigma_2} \right|, \]
  \[ 1 \leq \lambda. \]  
  \[ (112) \]

\[ (113) \]

• We call a multivariate tree (correlation-) optimal if it guarantees valid probabilities as long as
  
  \[ -1 + O(\sqrt{\Delta t}) \leq \rho \leq 1 - O(\sqrt{\Delta t}), \]

  such as the above one.\(^a\)

\(^a\)W. Kao (R98922093) (2011); W. Kao (R98922093), Lyuu, & Wen (D94922003) (2014).
Correlated Trinomial Model Simplified (continued)

- But this model cannot price 2-asset 2-barrier options accurately.\(^a\)

- Few multivariate trees are both optimal and able to handle multiple barriers.\(^b\)

- An alternative is to use orthogonalization.\(^c\)

\(^a\)See Y. Chang (B89704039, R93922034), Hsu (R7526001, D89922012), & Lyuu (2006); W. Kao (R98922093), Lyuu, & Wen (D94922003) (2014) for solutions.

\(^b\)See W. Kao (R98922093), Lyuu, & Wen (D94922003) (2014) for an exception.

\(^c\)Hull & White (1990); Dai (B82506025, R86526008, D8852600), C. Wang (F95922018), & Lyuu (2013).
Correlated Trinomial Model Simplified (concluded)

• Suppose we allow each asset’s volatility to be a function of time.\(^a\)

• There are \(k\) assets.

• Can you build an optimal multivariate tree that can handle two barriers on each asset in time \(O(n^{k+1})?\(^b\)

\(^a\)Recall p. 311.
\(^b\)See Y. Zhang (R05922052) (2019) for a complete solution.
Extrapolation

- It is a method to speed up numerical convergence.
- Say \( f(n) \) converges to an unknown limit \( f \) at rate of \( 1/n \):

\[
f(n) = f + \frac{c}{n} + o\left(\frac{1}{n}\right).
\]  

(114)

- Assume \( c \) is an unknown constant independent of \( n \).
  - Convergence is basically monotonic and smooth.
Extrapolation (concluded)

• From two approximations \( f(n_1) \) and \( f(n_2) \) and ignoring the smaller terms,
  \[
  f(n_1) = f + \frac{c}{n_1},
  \]
  \[
  f(n_2) = f + \frac{c}{n_2}.
  \]

• A better approximation to the desired \( f \) is
  \[
  f = \frac{n_1 f(n_1) - n_2 f(n_2)}{n_1 - n_2}.
  \]
  \begin{equation}
  (115)
  \end{equation}

• This estimate should converge faster than \( 1/n \).\(^a\)

• The Richardson extrapolation uses \( n_2 = 2n_1 \).

\(^a\)It is identical to the forward rate formula (22) on p. 150!
Improving BOPM with Extrapolation

- Consider standard European options.
- Denote the option value under BOPM using \( n \) time periods by \( f(n) \).
- It is known that BOPM convergences at the rate of \( 1/n \), consistent with Eq. (114) on p. 820.
- The plots on p. 302 (redrawn on next page) show that convergence to the true option value oscillates with \( n \).
- Extrapolation is inapplicable at this stage.
Improving BOPM with Extrapolation (concluded)

- Take the at-the-money option in the left plot on p. 823.
- The sequence with odd $n$ turns out to be monotonic and smooth (see the left plot on p. 825).\(^a\)
- Apply extrapolation (115) on p. 821 with $n_2 = n_1 + 2$, where $n_1$ is odd.
- Result is shown in the right plot on p. 825.
- The convergence rate is amazing.
- See Exercise 9.3.8 (p. 111) of the text for ideas in the general case.

\(^a\)This can be proved (L. Chang & Palmer, 2007).
Numerical Methods
All science is dominated by the idea of approximation.
— Bertrand Russell
Finite-Difference Methods

- Place a grid of points on the space over which the desired function takes value.
- Then approximate the function value at each of these points (p. 829).
- Solve the equation numerically by introducing difference equations in place of derivatives.
Example: Poisson’s Equation

- It is $\frac{\partial^2 \theta}{\partial x^2} + \frac{\partial^2 \theta}{\partial y^2} = -\rho(x, y)$, which describes the electrostatic field.

- Replace second derivatives with finite differences through central difference.

- Introduce evenly spaced grid points with distance of $\Delta x$ along the $x$ axis and $\Delta y$ along the $y$ axis.

- The finite-difference form is

$$-\rho(x_i, y_j) = \frac{\theta(x_{i+1}, y_j) - 2\theta(x_i, y_j) + \theta(x_{i-1}, y_j)}{(\Delta x)^2}$$

$$+ \frac{\theta(x_i, y_{j+1}) - 2\theta(x_i, y_j) + \theta(x_i, y_{j-1})}{(\Delta y)^2}.$$
Example: Poisson’s Equation (concluded)

- In the above, \( \Delta x \triangleq x_i - x_{i-1} \) and \( \Delta y \triangleq y_j - y_{j-1} \) for \( i, j = 1, 2, \ldots \).

- When the grid points are evenly spaced in both axes so that \( \Delta x = \Delta y = h \), the difference equation becomes

  \[
  -h^2 \rho(x_i, y_j) = \theta(x_{i+1}, y_j) + \theta(x_{i-1}, y_j) \\
  + \theta(x_i, y_{j+1}) + \theta(x_i, y_{j-1}) - 4\theta(x_i, y_j).
  \]

- Given boundary values, we can solve for the \( x_i \)s and the \( y_j \)s within the square \( [\pm L, \pm L] \).

- From now on, \( \theta_{i,j} \) will denote the finite-difference approximation to the exact \( \theta(x_i, y_j) \).
Explicit Methods

- Consider the diffusion equation
  \[ D(\frac{\partial^2 \theta}{\partial x^2}) - (\frac{\partial \theta}{\partial t}) = 0, \quad D > 0. \]

- Use evenly spaced grid points \((x_i, t_j)\) with distances \(\Delta x\) and \(\Delta t\), where \(\Delta x \overset{\Delta}{=} x_{i+1} - x_i\) and \(\Delta t \overset{\Delta}{=} t_{j+1} - t_j\).

- Employ central difference for the second derivative and forward difference for the time derivative to obtain

\[
\left. \frac{\partial \theta(x, t)}{\partial t} \right|_{t=t_j} = \frac{\theta(x, t_{j+1}) - \theta(x, t_j)}{\Delta t} + \cdots, \tag{116}
\]

\[
\left. \frac{\partial^2 \theta(x, t)}{\partial x^2} \right|_{x=x_i} = \frac{\theta(x_{i+1}, t) - 2\theta(x_i, t) + \theta(x_{i-1}, t)}{(\Delta x)^2} + \cdots \tag{117}
\]
Explicit Methods (continued)

- Next, assemble Eqs. (116) and (117) into a single equation at \((x_i, t_j)\).
- But we need to decide how to evaluate \(x\) in the first equation and \(t\) in the second.
- Since central difference around \(x_i\) is used in Eq. (117), we might as well use \(x_i\) for \(x\) in Eq. (116).
- Two choices are possible for \(t\) in Eq. (117).
- The first choice uses \(t = t_j\) to yield the following finite-difference equation,

\[
\frac{\theta_{i,j+1} - \theta_{i,j}}{\Delta t} = D \frac{\theta_{i+1,j} - 2\theta_{i,j} + \theta_{i-1,j}}{(\Delta x)^2}.
\] (118)
Explicit Methods (continued)

• The stencil of grid points involves four values, \( \theta_{i,j+1} \), \( \theta_{i,j} \), \( \theta_{i+1,j} \), and \( \theta_{i-1,j} \).

• Rearrange Eq. (118) on p. 833 as

\[
\theta_{i,j+1} = \frac{D \Delta t}{(\Delta x)^2} \theta_{i+1,j} + \left(1 - \frac{2D \Delta t}{(\Delta x)^2}\right) \theta_{i,j} + \frac{D \Delta t}{(\Delta x)^2} \theta_{i-1,j}.
\]

• We can calculate \( \theta_{i,j+1} \) from \( \theta_{i,j}, \theta_{i+1,j}, \theta_{i-1,j} \), at the previous time \( t_j \) (see exhibit (a) on next page).
Stencils

(a) $x_{i-1}$ $x_i$ $x_{i+1}$

(b) $x_{i-1}$ $x_i$ $x_{i+1}$

$t_j$ $t_{j+1}$
Explicit Methods (concluded)

• Starting from the initial conditions at \( t_0 \), that is, \( \theta_{i,0} = \theta(x_i, t_0), \ i = 1, 2, \ldots \), we calculate

\[
\theta_{i,1}, \quad i = 1, 2, \ldots.
\]

• And then

\[
\theta_{i,2}, \quad i = 1, 2, \ldots.
\]

• And so on.
Stability

• The explicit method is numerically unstable unless

\[ \Delta t \leq (\Delta x)^2 / (2D). \]

  – A numerical method is unstable if the solution is highly sensitive to changes in initial conditions.

• The stability condition may lead to high running times and memory requirements.

• For instance, halving \( \Delta x \) would imply quadrupling \( (\Delta t)^{-1} \), resulting in a running time 8 times as much.
Explicit Method and Trinomial Tree

• Recall that

\[ \theta_{i,j+1} = \frac{D\Delta t}{(\Delta x)^2} \theta_{i+1,j} + \left(1 - \frac{2D\Delta t}{(\Delta x)^2}\right) \theta_{i,j} + \frac{D\Delta t}{(\Delta x)^2} \theta_{i-1,j}. \]

• When the stability condition is satisfied, the three coefficients for \( \theta_{i+1,j}, \theta_{i,j}, \) and \( \theta_{i-1,j} \) all lie between zero and one and sum to one.

• They can be interpreted as probabilities.

• So the finite-difference equation becomes identical to backward induction on trinomial trees!
Explicit Method and Trinomial Tree (concluded)

- The freedom in choosing $\Delta x$ corresponds to similar freedom in the construction of trinomial trees.

- The explicit finite-difference equation is also identical to backward induction on a binomial tree.$^a$
  - Let the binomial tree take 2 steps each of length $\Delta t/2$.
  - It is now a trinomial tree.

---

$^a$Hilliard (2014).
Implicit Methods

- Suppose we use $t = t_{j+1}$ in Eq. (117) on p. 832 instead.
- The finite-difference equation becomes

$$\frac{\theta_{i,j+1} - \theta_{i,j}}{\Delta t} = D \frac{\theta_{i+1,j+1} - 2\theta_{i,j+1} + \theta_{i-1,j+1}}{(\Delta x)^2}.$$  \hfill (119)

- The stencil involves $\theta_{i,j}$, $\theta_{i,j+1}$, $\theta_{i+1,j+1}$, and $\theta_{i-1,j+1}$.
- This method is implicit:
  - The value of any one of the three quantities at $t_{j+1}$ cannot be calculated unless the other two are known.
  - See exhibit (b) on p. 835.
Implicit Methods (continued)

• Equation (119) can be rearranged as

\[ \theta_{i-1,j+1} - (2 + \gamma) \theta_{i,j+1} + \theta_{i+1,j+1} = -\gamma \theta_{i,j}, \]

where \( \gamma \equiv (\Delta x)^2 / (D\Delta t) \).

• This equation is unconditionally stable.

• Suppose the boundary conditions are given at \( x = x_0 \) and \( x = x_{N+1} \).

• After \( \theta_{i,j} \) has been calculated for \( i = 1, 2, \ldots, N \), the values of \( \theta_{i,j+1} \) at time \( t_{j+1} \) can be computed as the solution to the following tridiagonal linear system,
Implicit Methods (continued)

\[
\begin{bmatrix}
  a & 1 & 0 & \cdots & \cdots & \cdots & 0 \\
  1 & a & 1 & 0 & \cdots & \cdots & 0 \\
  0 & 1 & a & 1 & 0 & \cdots & 0 \\
  \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
  \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
  0 & \cdots & \cdots & 0 & 1 & a & 1 \\
  0 & \cdots & \cdots & \cdots & 0 & 1 & a \\
\end{bmatrix}
\begin{bmatrix}
  \theta_{1,j+1} \\
  \theta_{2,j+1} \\
  \theta_{3,j+1} \\
  \vdots \\
  \vdots \\
  \theta_{N,j+1} \\
\end{bmatrix}
= 
\begin{bmatrix}
  -\gamma \theta_{1,j} - \theta_{0,j+1} \\
  -\gamma \theta_{2,j} \\
  -\gamma \theta_{3,j} \\
  \vdots \\
  \vdots \\
  -\gamma \theta_{N-1,j} \\
  -\gamma \theta_{N,j} - \theta_{N+1,j+1} \\
\end{bmatrix},
\]

where \( a \triangleq -2 - \gamma \).
Implicit Methods (concluded)

- Tridiagonal systems can be solved in $O(N)$ time and $O(N)$ space.
  - Never invert a matrix to solve a tridiagonal system.
- The matrix above is nonsingular when $\gamma \geq 0$.
  - A square matrix is nonsingular if its inverse exists.
Crank-Nicolson Method

• Take the average of explicit method (118) on p. 833 and implicit method (119) on p. 840:

\[
\frac{\theta_{i,j+1} - \theta_{i,j}}{\Delta t} = \frac{1}{2} \left( \frac{D \theta_{i+1,j} - 2\theta_{i,j} + \theta_{i-1,j}}{(\Delta x)^2} + D \frac{\theta_{i+1,j} + 1 - 2\theta_{i,j} + 1 + \theta_{i-1,j} + 1}{(\Delta x)^2} \right).
\]

• After rearrangement,

\[
\gamma \theta_{i,j+1} - \frac{\theta_{i+1,j} + 1 - 2\theta_{i,j} + 1 + \theta_{i-1,j} + 1}{2} = \gamma \theta_{i,j} + \frac{\theta_{i+1,j} - 2\theta_{i,j} + \theta_{i-1,j}}{2}.
\]

• This is an unconditionally stable implicit method with excellent rates of convergence.
Stencil

\[ x_{i+1} \]
\[ x_i \]
\[ x_{i+1} \]
\[ t_j \]
\[ t_{j+1} \]
Numerically Solving the Black-Scholes PDE (90) on p. 678

• See text.

• Brennan and Schwartz (1978) analyze the stability of the implicit method.
Monte Carlo Simulation\textsuperscript{a}

- Monte Carlo simulation is a sampling scheme.
- In many important applications within finance and without, Monte Carlo is one of the few feasible tools.
- When the time evolution of a stochastic process is not easy to describe analytically, Monte Carlo may very well be the only strategy that succeeds consistently.

\textsuperscript{a}A top 10 algorithm (Dongarra & Sullivan, 2000).
The Big Idea

- Assume $X_1, X_2, \ldots, X_n$ have a joint distribution.

- $\theta \triangleq E[g(X_1, X_2, \ldots, X_n)]$ for some function $g$ is desired.

- We generate

  \[
  \left( x_1^{(i)}, x_2^{(i)}, \ldots, x_n^{(i)} \right), \quad 1 \leq i \leq N
  \]

  independently with the same joint distribution as $(X_1, X_2, \ldots, X_n)$.

- Set

  \[
  Y_i \triangleq g \left( x_1^{(i)}, x_2^{(i)}, \ldots, x_n^{(i)} \right).
  \]
The Big Idea (concluded)

- $Y_1, Y_2, \ldots, Y_N$ are independent and identically distributed random variables.
- Each $Y_i$ has the same distribution as $Y \overset{\Delta}{=} g(X_1, X_2, \ldots, X_n)$.
- Since the average of these $N$ random variables, $\bar{Y}$, satisfies $E[\bar{Y}] = \theta$, it can be used to estimate $\theta$.
- The strong law of large numbers says that this procedure converges almost surely.
- The number of replications (or independent trials), $N$, is called the sample size.
Accuracy

• The Monte Carlo estimate and true value may differ owing to two reasons:
  1. Sampling variation.
  2. The discreteness of the sample paths.\(^a\)

• The first can be controlled by the number of replications.

• The second can be controlled by the number of observations along the sample path.

\(^a\)This may not be an issue if the financial derivative only requires discrete sampling along the time dimension, such as the discrete barrier option.
Accuracy and Number of Replications

- The statistical error of the sample mean $\overline{Y}$ of the random variable $Y$ grows as $1/\sqrt{N}$.
  - Because $\text{Var}[\overline{Y}] = \text{Var}[Y]/N$.

- In fact, this convergence rate is asymptotically optimal.\(^{a}\)

- So the variance of the estimator $\overline{Y}$ can be reduced by a factor of $1/N$ by doing $N$ times as much work.

- This is amazing because the same order of convergence holds independently of the dimension $n$.

\(^{a}\)The Berry-Esseen theorem.
Accuracy and Number of Replications (concluded)

• In contrast, classic numerical integration schemes have an error bound of $O(N^{-c/n})$ for some constant $c > 0$. 
  – $n$ is the dimension.

• The required number of evaluations thus grows exponentially in $n$ to achieve a given level of accuracy. 
  – The curse of dimensionality.

• The Monte Carlo method is more efficient than alternative procedures for multivariate derivatives when $n$ is large.
Monte Carlo Option Pricing

- For the pricing of European options on a dividend-paying stock, we may proceed as follows.
- Assume
  \[ \frac{dS}{S} = \mu \, dt + \sigma \, dW. \]
- Stock prices \( S_1, S_2, S_3, \ldots \) at times \( \Delta t, 2\Delta t, 3\Delta t, \ldots \) can be generated via
  \[
  S_{i+1} = S_i e^{(\mu - \frac{\sigma^2}{2}) \Delta t + \sigma \sqrt{\Delta t} \, \xi}, \quad \xi \sim \mathcal{N}(0, 1), \quad (120)
  \]
  by Eq. (84) on p. 616.
Monte Carlo Option Pricing (continued)

- If we discretize \( dS/S = \mu \, dt + \sigma \, dW \) directly, we will obtain

\[
S_{i+1} = S_i + S_i \mu \, \Delta t + S_i \sigma \sqrt{\Delta t} \, \xi.
\]

- But this is locally normally distributed, not lognormally, hence biased.

- In practice, this is not expected to be a major problem as long as \( \Delta t \) is sufficiently small.

---

\(^a\)Contributed by Mr. Tai, Hui-Chin (R97723028) on April 22, 2009.
Monte Carlo Option Pricing (continued)

Non-dividend-paying stock prices in a risk-neutral economy can be generated by setting $\mu = r$ and $\Delta t = T$.

1: $C := 0$; \{Accumulated terminal option value.\}
2: \textbf{for} $i = 1, 2, 3, \ldots , N$ \textbf{do}
3: \hspace{1em} $P := S \times e^{(r-\sigma^2/2)T+\sigma\sqrt{T}\xi}$, $\xi \sim N(0, 1)$;
4: \hspace{1em} $C := C + \max(P - X, 0)$;
5: \hspace{1em} \textbf{end for}
6: \textbf{return} $Ce^{-rT}/N$;
Monte Carlo Option Pricing (concluded)

Pricing Asian options is also easy.

1: $C := 0$
2: for $i = 1, 2, 3, \ldots , N$ do
3: $P := S; M := S$
4: for $j = 1, 2, 3, \ldots , n$ do
5: $P := P \times e^{(r-\sigma^2/2)(T/n)+\sigma\sqrt{T/n} \xi}$;
6: $M := M + P$
7: end for
8: $C := C + \max(M/(n+1) - X, 0)$;
9: end for
10: return $Ce^{-rT}/N$;
How about American Options?

- Standard Monte Carlo simulation is inappropriate for American options because of early exercise.
  - Given a sample path $S_0, S_1, \ldots, S_n$, how to decide which $S_i$ is an early-exercise point?
  - What is the option price at each $S_i$ if the option is not exercised?

- It is difficult to determine the early-exercise point based on one single path.

- But Monte Carlo simulation can be modified to price American options with small biases (pp. 919ff).\(^a\)

\(^a\)Longstaff & Schwartz (2001).
Delta and Common Random Numbers

• In estimating delta, it is natural to start with the finite-difference estimate

\[ e^{-r\tau} \frac{E[P(S + \epsilon)] - E[P(S - \epsilon)]}{2\epsilon} . \]

– \( P(x) \) is the terminal payoff of the derivative security when the underlying asset’s initial price equals \( x \).

• Use simulation to estimate \( E[P(S + \epsilon)] \) first.
• Use another simulation to estimate \( E[P(S - \epsilon)] \).
• Finally, apply the formula to approximate the delta.
• This is also called the bump-and-revalue method.
Delta and Common Random Numbers (concluded)

• This method is not recommended because of its high variance.

• A much better approach is to use common random numbers to lower the variance:

\[ e^{-r\tau} E \left[ \frac{P(S + \epsilon) - P(S - \epsilon)}{2\epsilon} \right] . \]

• Here, the same random numbers are used for \( P(S + \epsilon) \) and \( P(S - \epsilon) \).

• This holds for gamma and cross gamma.\(^a\)

\(^a\)For multivariate derivatives.
Problems with the Bump-and-Revalue Method

• Consider the binary option with payoff

\[
\begin{cases}
1, & \text{if } S(T) > X, \\
0, & \text{otherwise}.
\end{cases}
\]

• Then

\[
P(S+\epsilon)-P(S-\epsilon) = \begin{cases}
1, & \text{if } S + \epsilon > X \text{ and } S - \epsilon < X, \\
0, & \text{otherwise}.
\end{cases}
\]

• So the finite-difference estimate per run for the (undiscounted) delta is 0 or \(O(1/\epsilon)\).

• This means high variance.
Problems with the Bump-and-Revalue Method (concluded)

- The price of the binary option equals
  \[ e^{-r\tau} N(x - \sigma \sqrt{\tau}). \]
  - It equals \textit{minus} the derivative of the European call with respect to \( X \).
  - It also equals \( X\tau \) times the rho of a European call (p. 358).

- Its delta is
  \[ \frac{N'(x - \sigma \sqrt{\tau})}{S\sigma \sqrt{\tau}}. \]
Gamma

- The finite-difference formula for gamma is

\[ e^{-r\tau} E \left[ \frac{P(S + \epsilon) - 2 \times P(S) + P(S - \epsilon)}{\epsilon^2} \right] . \]

- For a correlation option with multiple underlying assets, the finite-difference formula for the cross gamma

\[ \partial^2 P(S_1, S_2, \ldots) / (\partial S_1 \partial S_2) \] is:

\[ e^{-r\tau} E \left[ \frac{P(S_1 + \epsilon_1, S_2 + \epsilon_2) - P(S_1 - \epsilon_1, S_2 + \epsilon_2)}{4\epsilon_1\epsilon_2} \right. \]

\[ \left. - P(S_1 + \epsilon_1, S_2 - \epsilon_2) + P(S_1 - \epsilon_1, S_2 - \epsilon_2) \right] . \]
Gamma (continued)

• Choosing an \( \epsilon \) of the right magnitude can be challenging.
  – If \( \epsilon \) is too large, inaccurate Greeks result.
  – If \( \epsilon \) is too small, unstable Greeks result.

• This phenomenon is sometimes called the curse of differentiation.\(^a\)

\(^a\)Aït-Sahalia & Lo (1998); Bondarenko (2003).
Gamma (continued)

• In general, suppose (in some sense)

\[
\frac{\partial^i}{\partial \theta^i} e^{-r\tau} E[P(S)] = e^{-r\tau} E \left[ \frac{\partial^i P(S)}{\partial \theta^i} \right]
\]

holds for all \( i > 0 \), where \( \theta \) is a parameter of interest.a

  – A common requirement is Lipschitz continuity.b

• Then Greeks become integrals.

• As a result, we avoid \( \epsilon \), finite differences, and resimulation.

---

\[ a \] \( \partial^i P(S)/\partial \theta^i \) may not be partial differentiation in the classic sense.

\[ b \] Broadie & Glasserman (1996).
Gamma (continued)

• This is indeed possible for a broad class of payoff functions.\(^a\)
  
  – Roughly speaking, any payoff function that is equal to a sum of products of differentiable functions and indicator functions with the right kind of support.
  
  – For example, the payoff of a call is

\[
\max(S(T) - X, 0) = (S(T) - X)I\{S(T) - X \geq 0\}.
\]

  – The results are too technical to cover here (see next page).

\(^a\)Teng (R91723054) (2004); Lyuu & Teng (R91723054) (2011).
Gamma (continued)

• Suppose \( h(\theta, x) \in H \) with pdf \( f(x) \) for \( x \) and \( g_j(\theta, x) \in G \) for \( j \in \mathcal{B} \), a finite set of natural numbers.

• Then

\[
\frac{\partial}{\partial \theta} \int_{\mathbb{R}} h(\theta, x) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\theta, x) > 0\}}(x) f(x) \, dx \n\]

\[
= \int_{\mathbb{R}} h_\theta(\theta, x) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\theta, x) > 0\}}(x) f(x) \, dx \n\]

\[
+ \sum_{l \in \mathcal{B}} \left[ h(\theta, x) J_l(\theta, x) \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_j(\theta, x) > 0\}}(x) f(x) \right]_{x = \chi_l(\theta)}, \n\]

where

\[
J_l(\theta, x) = \text{sign} \left( \frac{\partial g_l(\theta, x)}{\partial x_k} \right) \frac{\partial g_l(\theta, x) / \partial \theta}{\partial g_l(\theta, x) / \partial x} \text{ for } l \in \mathcal{B}. \n\]
Gamma (concluded)

- Similar results have been derived for Levy processes.\textsuperscript{a}
- Formulas are also recently obtained for credit derivatives.\textsuperscript{b}
- In queueing networks, this is called infinitesimal perturbation analysis (IPA).\textsuperscript{c}

\textsuperscript{a}Lyuu, Teng (R91723054), & S. Wang (2013).
\textsuperscript{b}Lyuu, Teng (R91723054), & Tseng (2014, 2018).
\textsuperscript{c}Cao (1985); Y. C. Ho & Cao (1985).