American Option Pricing by Simulation

- The continuation value of an American option is the conditional expectation of the payoff from keeping the option alive now.
- The option holder must compare the immediate exercise value and the continuation value.
- In standard Monte Carlo simulation, each path is treated independently of other paths.
- But the exercise decision cannot be reached by looking at just one path.

The Least-Squares Monte Carlo Approach

- Estimate the continuation value from the cross-sectional information in the simulation with least squares.^a
- The result is a function of the state for estimating it.
- Use the estimated continuation value for each path to determine its cash flow.
- This is called least-squares Monte Carlo (LSM).

^aLongstaff & Schwartz (2001).

The Least-Squares Monte Carlo Approach (concluded)

- LSM is provably convergent.^a
- LSM can be easily parallelized.^b
 - Partition the paths into subproblems and perform LSM on each independently.
 - The speedup is close to linear (i.e., proportional to the number of cores).
- Surprisingly, accuracy is not affected.

^aClément, Lamberton, & Protter (2002); Stentoft (2004). ^bK. Huang (B96902079, R00922018) (2013); C. W. Chen (B97902046, R01922005) (2014); C. W. Chen (B97902046, R01922005), K. Huang (B96902079, R00922018) & Lyuu (2015).

A Numerical Example

- Consider a 3-year American put on a non-dividend-paying stock.
- The put is exercisable at years 0, 1, 2, and 3.
- The strike price X = 105.
- The annualized riskless rate is r = 5%.
 - The annual discount factor equals 0.951229.
- The current stock price is 101.
- We use 8 price paths to illustrate the algorithm.

		Stock price	e paths	
Path	Year 0	Year 1	Year 2	Year 3
1	101	97.6424	92.5815	107.5178
2	101	101.2103	105.1763	102.4524
3	101	105.7802	103.6010	124.5115
4	101	96.4411	98.7120	108.3600
5	101	124.2345	101.0564	104.5315
6	101	95.8375	93.7270	99.3788
7	101	108.9554	102.4177	100.9225
8	101	104.1475	113.2516	115.0994



- We use the basis functions $1, x, x^2$.
 - Other basis functions are possible.^a
- The plot next page shows the final estimated optimal exercise strategy given by LSM.
- We now proceed to tackle our problem.
- The idea is to calculate the cash flow along each path, using information from *all* in-the-money paths.

^aLaguerre polynomials, Hermite polynomials, Legendre polynomials, Chebyshev polynomials, Gedenbauer polynomials, or Jacobi polynomials.



A Numerical Example (continued)					
	Cash	n flows at	year 3		
Path	Year 0	Year 1	Year 2	Year 3	
1				0	
2				2.5476	
3				0	
4				0	
5				0.4685	
6				5.6212	
7				4.0775	
8				0	

- The cash flows at year 3 are the put's payoffs.
- Only 4 paths are in the money: 2, 5, 6, 7.
- Some of the cash flows may not materialize if the put is exercised earlier, which we will find out later.
- Incidentally, the *European* counterpart has a value of

$$0.951229^3 \times \frac{2.5476 + 0.4685 + 5.6212 + 4.0775}{8}$$

= 1.3680.

- We move on to year 2.
- For each state that is in the money at year 2, we must decide whether to exercise it.
- There are 6 paths for which the put is in the money: 1, 3, 4, 5, 6, 7.^a
- Only in-the-money paths will be used in the regression because they are where early exercise is possible.

- If there were none, move on to year 1.

^aRecall p. 931.

- Let x denote the stock price at year 2 for each of those 6 paths.
- Let y denote the corresponding discounted future cash flow (at year 3) if the put is *not* exercised at year 2.

	Regressic	on at year 2
Path	x	y
1	92.5815	0 imes 0.951229
2		
3	103.6010	0 imes 0.951229
4	98.7120	0 imes 0.951229
5	101.0564	0.4685 imes 0.951229
6	93.7270	5.6212 imes 0.951229
7	102.4177	4.0775×0.951229
8		

- We regress y on 1, x, and x^2 .
- The result is

 $f(x) = 22.08 - 0.313114 \times x + 0.00106918 \times x^2.$

- f(x) estimates the *continuation value* conditional on the stock price at year 2.
- We next compare the immediate exercise value and the estimated continuation value.^a

^aThe f(102.4177) entry on the next page was corrected by Mr. Tu, Yung-Szu (B79503054, R83503086) on May 25, 2017.

ercise decision at year 2	al early exe	Optim
Continuation	Exercise	Path
f(92.5815) = 2.2558	12.4185	1
		2
f(103.6010) = 1.1168	1.3990	3
f(98.7120) = 1.5901	6.2880	4
f(101.0564) = 1.3568	3.9436	5
f(93.7270) = 2.1253	11.2730	6
f(102.4177) = 1.2266	2.5823	7
		8

O2023 Prof. Yuh-Dauh Lyuu, National Taiwan University

- The put should be exercised in all 6 paths: 1, 3, 4, 5, 6,
 7.
- Now, any positive cash flow at year 3 vanishes for these paths as the put has been exercised before it.^a

- They are paths 5, 6, 7.

• The cash flows on p. 935 become the ones on next slide.

^aRecall p. 931.

	A Numerical Example (continued)				
		Cash f	lows at ye	ears 2 & 3	
	Path	Year 0	Year 1	Year 2	Year 3
_	1			12.4185	0
	2			0	2.5476
	3			1.3990	0
	4			6.2880	0
	5			3.9436	0
	6			11.2730	0
	7			2.5823	0
	8			0	0
_					

- We move on to year 1.
- For each state that is in the money at year 1, we must decide whether to exercise it.
- There are 5 paths for which the put is in the money: 1, 2, 4, 6, 8.^a
- Only in-the-money paths will be used in the regression because they are where early exercise is possible.

- If there were none, move on to year 0.

^aRecall p. 931.

- Let x denote the stock price at year 1 for each of those 5 paths.
- Let y denote the corresponding discounted future cash flow if the put is not exercised at year 1.
- From p. 943, we have the following table.

A Numerical Example (continued)					
on at year 1	Regressio				
y	x	Path			
12.4185×0.951229	97.6424	1			
2.5476×0.951229^2	101.2103	2			
		3			
6.2880 imes 0.951229	96.4411	4			
		5			
11.2730×0.951229	95.8375	6			
		7			
0×0.951229	104.1475	8			

- We regress y on 1, x, and x^2 .
- The result is

 $f(x) = -420.964 + 9.78113 \times x - 0.0551567 \times x^2.$

- f(x) estimates the continuation value conditional on the stock price at year 1.
- We next compare the immediate exercise value and the estimated continuation value.

-		
Continuation	Exercise	Path
f(97.6424) = 8.2230	7.3576	1
f(101.2103) = 3.9882	3.7897	2
		3
f(96.4411) = 9.3329	8.5589	4
		5
f(95.8375) = 9.83042	9.1625	6
		7
f(104.1475) = -0.551885	0.8525	8

Optimal early exercise decision at year 1

• The put should be exercised for 1 path only: 8.

- Note that its f(104.1475) < 0.

- Now, any positive future cash flow vanishes for this path.
 But there is none.
- The cash flows on p. 943 become the ones on next slide.
- They also confirm the plot on p. 934.

A Numerical Example (continued)				
	Cash flo	ws at yea	rs 1, 2, & 3	3
Path	Year 0	Year 1	Year 2	Year 3
1		0	12.4185	0
2		0	0	2.5476
3		0	1.3990	0
4		0	6.2880	0
5		0	3.9436	0
6		0	11.2730	0
7		0	2.5823	0
8		0.8525	0	0

- We move on to year 0.
- The continuation value is, from p 950,

 $(12.4185 \times 0.951229^{2} + 2.5476 \times 0.951229^{3} + 1.3990 \times 0.951229^{2} + 6.2880 \times 0.951229^{2} + 3.9436 \times 0.951229^{2} + 11.2730 \times 0.951229^{2} + 2.5823 \times 0.951229^{2} + 0.8525 \times 0.951229)/8$

= 4.66263.

• As this is larger than the immediate exercise value of

105 - 101 = 4,

the put should not be exercised at year 0.

- Hence the put's value is estimated to be 4.66263.
- Compare this with the European put's value of 1.3680.^a

^aRecall p. 936.

Time Series Analysis

The historian is a prophet in reverse. — Friedrich von Schlegel (1772–1829)

Even in my tape reading something enters that is more than mere arithmetic.
— Edwin Lefèvre (1971–1943), Reminiscences of a Stock Operator (1923)

GARCH Option Pricing

- Options can be priced when the underlying asset's return follows a GARCH (generalized autoregressive conditional heteroskedastic) process.^a
- 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 .

^aBollerslev (1986) and Taylor (1986). They are the "most popular models for time-varying volatility" (Alexander, 2001). A Bloomberg quant said to me on Feb 29, 2008, that GARCH is seldom used in trading.

• Adopt the following risk-neutral process for price:^a

$$\ln \frac{S_{t+1}}{S_t} = r - \frac{h_t^2}{2} + h_t \epsilon_{t+1}, \qquad (127)$$

where

$$h_{t+1}^2 = \beta_0 + \beta_1 h_t^2 + \beta_2 h_t^2 (\epsilon_{t+1} - c)^2, \qquad (128)$$

 $\epsilon_{t+1} \sim N(0,1)$ given information at date t,

r = daily riskless return,

$$c \geq 0.$$

• This is called the nonlinear asymmetric GARCH (or NGARCH) model.

^aDuan (1995).

- The five unknown parameters of the model are c, h_0, β_0, β_1 , and β_2 .
- It is postulated that $\beta_0, \beta_1, \beta_2 \ge 0$ to make the conditional variance positive.
- There are other inequalities to satisfy such as $\beta_1 + \beta_2 < 1$ (see text).
- It can be shown that $h_t^2 \ge \min \left[h_0^2, \beta_0 / (1 \beta_1) \right]$.^a

^aLyuu & C. Wu (**R90723065**) (2005).

- It captures the volatility clustering in asset returns first noted by Mandelbrot (1963).^a
 - When c = 0, a large ϵ_{t+1} results in a large h_{t+1} , which in turns tends to yield a large h_{t+2} , and so on.
- It also captures the negative correlation between the asset return and changes in its (conditional) volatility.^b
 - For c > 0, a positive ϵ_{t+1} (good news) tends to decrease h_{t+1} , whereas a negative ϵ_{t+1} (bad news) tends to do the opposite.

^a"... large changes tend to be followed by large changes—of either sign—and small changes tend to be followed by small changes"

^bNoted by Black (1976): Volatility tends to rise in response to "bad news" and fall in response to "good news."

- This is called the leverage effect.
 - A falling stock price raises the fixed costs, relatively speaking.^a
 - Thus c is called the leverage effect parameter.
- With $y_t \stackrel{\Delta}{=} \ln S_t$ denoting the logarithmic price, the model becomes

$$y_{t+1} = y_t + r - \frac{h_t^2}{2} + h_t \epsilon_{t+1}.$$
 (129)

• The pair (y_t, h_t^2) completely describes the current state.

^aBlack (1992).

• The conditional mean and variance of y_{t+1} are clearly

$$E[y_{t+1} | y_t, h_t^2] = y_t + r - \frac{h_t^2}{2}, \qquad (130)$$

Var $[y_{t+1} | y_t, h_t^2] = h_t^2. \qquad (131)$

• Finally, given (y_t, h_t^2) , the correlation between y_{t+1} and h_{t+1} equals

$$-\frac{2c}{\sqrt{2+4c^2}},$$

which is negative for c > 0.

GARCH Model: Inferences

- Suppose the parameters c, h_0, β_0, β_1 , and β_2 are given.
- Then we can recover h_1, h_2, \ldots, h_n and $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ from the prices

$$S_0, S_1, \ldots, S_n$$

under the GARCH model (127) on p. 956.

• This is useful in statistical inferences.

The Ritchken-Trevor (RT) Algorithm $^{\rm a}$

- 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.^b
- We need to mitigate this combinatorial explosion.

```
<sup>a</sup>Ritchken & Trevor (1999).
<sup>b</sup>Why?
```
- Partition a day into n periods.
- Three states follow each state (y_t, h_t^2) after a period.
- As the trinomial model combines, each state at date t is followed by 2n + 1 states at date t + 1.^a
- These 2n + 1 values must approximate the distribution of (y_{t+1}, h_{t+1}^2) to guarantee convergence.
- So the conditional moments (130)-(131) at date t+1 on p. 960 must be matched by the trinomial model.

^aRecall p. 741.

- It remains to pick the jump size and the three branching probabilities.
- The role of σ in the Black-Scholes option pricing model is played by h_t in the GARCH model.
- As a jump size proportional to σ/\sqrt{n} is picked in the BOPM, a comparable magnitude will be chosen here.
- Define $\gamma \stackrel{\Delta}{=} h_0$, though other multiples of h_0 are possible.
- Let

$$\gamma_n \stackrel{\Delta}{=} \frac{\gamma}{\sqrt{n}}.$$

- The jump size will be some integer multiple η of γ_n .
- We call η the jump parameter (see next page).
- Clearly, the magnitude of η tends to grow with h_t .
- The middle branch does not change the underlying asset's price.



• The probabilities for the up, middle, and down branches are

$$p_u = \frac{h_t^2}{2\eta^2 \gamma^2} + \frac{r - (h_t^2/2)}{2\eta\gamma\sqrt{n}},$$
 (132)

$$p_m = 1 - \frac{h_t^2}{\eta^2 \gamma^2},$$
 (133)

$$p_d = \frac{h_t^2}{2\eta^2 \gamma^2} - \frac{r - (h_t^2/2)}{2\eta\gamma\sqrt{n}}.$$
 (134)

- It can be shown that:
 - The trinomial model takes on 2n + 1 values at date t + 1 for y_{t+1} .
 - These values match y_{t+1} 's mean.
 - These values match y_{t+1} 's variance asymptotically.
- The central limit theorem guarantees convergence to the continuous-space model as n increases.^a

^aAssume the probabilities are valid.

- We can dispense with the intermediate nodes between dates to create a (2n + 1)-nomial tree.^a
- The resulting model is multinomial with 2n + 1branches 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 larger.^b

^aSee p. 970.

^bContrast it with the case on p. 412.



• 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 , where

$$-n \leq \ell \leq n.$$

- To reach that price in n periods, the number of up moves must exceed that of down moves by exactly ℓ .
- The probability this happens is

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

with
$$j_u, j_m, j_d \ge 0, \ n = j_u + j_m + j_d$$
, and $\ell = j_u - j_d$.

• A simple way to calculate the $P(\ell)$ s starts by noting^a

$$(p_u x + p_m + p_d x^{-1})^n = \sum_{\ell=-n}^n P(\ell) x^{\ell}.$$
(135)

- Convince yourself that the "accounting" is done correctly.
- So we expand $(p_u x + p_m + p_d x^{-1})^n$ and retrieve the probabilities by reading off the coefficients.
- It can be computed in $O(n^2)$ time, if not less.

^aC. Wu (**R90723065**) (2003); Lyuu & C. Wu (**R90723065**) (2003, 2005).

- The updating rule (128) on p. 956 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) is associated with this variance:

$$h_{t+1}^2 = \beta_0 + \beta_1 h_t^2 + \beta_2 h_t^2 (\epsilon_{t+1}' - c)^2, \qquad (136)$$

– Above, the z-score^a

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

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

^aNote that the mean of ϵ'_{t+1} is $r - (h_t^2/2)$.

- Different h_t^2 may require different η so that the probabilities (132)–(134) on p. 967 lie between 0 and 1.
- This implies varying jump sizes $\eta \gamma_n$.
- The necessary requirement $p_m \ge 0$ implies $\eta \ge h_t/\gamma$.
- Hence we try

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

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

• The sufficient and necessary condition for valid probabilities to exist is^a

$$\frac{|r - (h_t^2/2)|}{2\eta\gamma\sqrt{n}} \le \frac{h_t^2}{2\eta^2\gamma^2} \le \min\left(1 - \frac{|r - (h_t^2/2)|}{2\eta\gamma\sqrt{n}}, \frac{1}{2}\right)$$

- The plot on p. 976 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 $\eta = 2$.

^aC. Wu (R90723065) (2003); Lyuu & C. Wu (R90723065) (2003, 2005).



- The topology of the tree is not a standard combining multinomial tree.
- For example, a few nodes on p. 976 such as nodes (2,0) and (2,-1) have *multiple* jump sizes.
- The reason is path dependency of the model.
 - Two paths can reach node (2,0) from the root node, each with a different variance h_t^2 for the node.
 - One variance results in $\eta = 1$.
 - The other results in $\eta = 2$.

The RT Algorithm (concluded)

- The number of possible values of h_t^2 at a node can be exponential.
 - Because each path may result in a different h_t^2 .
- To address this problem, we record only the maximum and minimum h_t^2 at each node.^a
- Therefore, each node on the tree contains only two states (y_t, h_{max}^2) and (y_t, h_{min}^2) .
- Each of (y_t, h_{\max}^2) and (y_t, h_{\min}^2) carries its own η and set of 2n + 1 branching probabilities.

^aCakici & Topyan (2000). But see p. 1013 for a potential problem.

Negative Aspects of the Ritchken-Trevor Algorithm $^{\rm a}$

- 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 quite limited in practice.
- The RT algorithm can be modified to be free of shortened maturity and exponential complexity.^b

^aLyuu & C. Wu (R90723065) (2003, 2005). ^bIts size is only $O(T^2)$ if $n \leq (\sqrt{(1-\beta_1)/\beta_2} - c)^2$, where T is the number of days to maturity!

Numerical Examples



• 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, $h^2(0,0) = h_0^2 = 0.0001096$.

- Let $h_{\max}^2(i,j)$ denote the maximum variance at node (i,j).
- Let $h_{\min}^2(i, j)$ denote the minimum variance at node (i, j).
- Initially, $h_{\max}^2(0,0) = h_{\min}^2(0,0) = h_0^2$.
- The resulting 3-day tree is depicted on p. 983.



- A top number inside a gray box refers to the minimum variance h_{\min}^2 for the node.
- A bottom number inside a gray box refers to the maximum variance h_{max}^2 for the node.
- Variances are multiplied by 100,000 for readability.
- The top number inside a white box refers to the η for h_{\min}^2 .
- The bottom number inside a white box refers to the η for h_{max}^2 .

- Let us see how the numbers are calculated.
- Start with the root node, node (0,0).
- Try $\eta = 1$ in Eqs. (132)–(134) on p. 967 first to obtain

 $p_u = 0.4974,$ $p_m = 0,$ $p_d = 0.5026.$

• As they are valid, the three branches from the root node take single jumps.

- Move on to node (1,1).
- It has one predecessor node—node (0,0)—and it takes an up move to reach node (1,1).
- So apply updating rule (136) on p. 973 with $\ell = 1$ and $h_t^2 = h^2(0,0)$.
- The result is $h^2(1,1) = 0.000109645$.

• Because $\lceil h(1,1)/\gamma \rceil = 2$, we try $\eta = 2$ in Eqs. (132)–(134) on p. 967 first to obtain

$$p_u = 0.1237,$$

 $p_m = 0.7499,$
 $p_d = 0.1264.$

• As they are valid, the three branches from node (1,1) take double jumps.

- Carry out similar calculations for node (1,0) with $\ell = 0$ in updating rule (136) on p. 973.
- Carry out similar calculations for node (1, -1) with $\ell = -1$ in updating rule (136).
- Single jump $\eta = 1$ works for both nodes.
- The resulting variances are

 $h^2(1,0) = 0.000105215,$ $h^2(1,-1) = 0.000109553.$

- 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 node (2,0), we apply updating rule (136) on p. 973 with $\ell = 0$ and $h_t^2 = h^2(1,0)$.
- The result is $h_{t+1}^2 = 0.000101269$.

- Now move on to the other predecessor node (1, -1).
- Because it takes an up move to reach node (2,0), apply updating rule (136) on p. 973 with $\ell = 1$ and $h_t^2 = h^2(1,-1)$.
- The result is $h_{t+1}^2 = 0.000109603$.
- We hence record

$$h_{\min}^2(2,0) = 0.000101269,$$

 $h_{\max}^2(2,0) = 0.000109603.$

- Consider state $h_{\max}^2(2,0)$ first.
- Because $\lceil h_{\max}(2,0)/\gamma \rceil = 2$, we first try $\eta = 2$ in Eqs. (132)–(134) on p. 967 to obtain

$$p_u = 0.1237,$$

 $p_m = 0.7500,$
 $p_d = 0.1263.$

• As they are valid, the three branches from node (2,0) with the maximum variance take double jumps.

- Now consider state $h_{\min}^2(2,0)$.
- Because $\lceil h_{\min}(2,0)/\gamma \rceil = 1$, we first try $\eta = 1$ in Eqs. (132)–(134) on p. 967 to obtain

 $p_u = 0.4596,$ $p_m = 0.0760,$ $p_d = 0.4644.$

• As they are valid, the three branches from node (2,0) with the minimum variance take single jumps.

- Node (2, -1) has 3 predecessor nodes.
- Start with node (1,1).
- Because it takes one down move to reach node (2, -1), we apply updating rule (136) on p. 973 with $\ell = -1$ and $h_t^2 = h^2(1, 1)$.^a
- The result is $h_{t+1}^2 = 0.0001227$.

^aNote that it is not $\ell = -2$. The reason is that h(1,1) has $\eta = 2$ (p. 987).

- Now move on to predecessor node (1, 0).
- Because it also takes a down move to reach node (2, -1), we apply updating rule (136) on p. 973 with $\ell = -1$ and $h_t^2 = h^2(1, 0)$.
- The result is $h_{t+1}^2 = 0.000105609$.

- Finally, consider predecessor node (1, -1).
- Because it takes a middle move to reach node (2, -1), we apply updating rule (136) on p. 973 with $\ell = 0$ and $h_t^2 = h^2(1, -1)$.
- The result is $h_{t+1}^2 = 0.000105173$.
- We hence record

$$h_{\min}^2(2,-1) = 0.000105173,$$

 $h_{\max}^2(2,-1) = 0.0001227.$

- Consider state $h_{\max}^2(2,-1)$.
- Because $\lceil h_{\max}(2,-1)/\gamma \rceil = 2$, we first try $\eta = 2$ in Eqs. (132)–(134) on p. 967 to obtain

$$p_u = 0.1385,$$

 $p_m = 0.7201,$
 $p_d = 0.1414.$

As they are valid, the three branches from node (2, −1) with the maximum variance take double jumps.

- Next, consider state $h_{\min}^2(2,-1)$.
- Because $\lceil h_{\min}(2,-1)/\gamma \rceil = 1$, we first try $\eta = 1$ in Eqs. (132)–(134) on p. 967 to obtain

 $p_u = 0.4773,$ $p_m = 0.0404,$ $p_d = 0.4823.$

As they are valid, the three branches from node (2, -1) with the minimum variance take 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 up to 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 $^{\rm a}$

- Recall the problems mentioned on p. 979.
- In our case, combinatorial explosion occurs when

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

(see the next plot).

- 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!

^aLyuu & C. Wu (**R90723065**) (2003, 2005).



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_{\max}^2 and h_{\min}^2 .
- We now increase that number to K equally spaced variances between h_{max}^2 and h_{min}^2 at each node.
- Besides the minimum and maximum variances, the other K-2 variances in between are linearly interpolated.^a

^aLog-linear interpolation works better in practice (Lyuu & C. Wu (R90723065), 2005). Log-cubic interpolation works even better (C. Liu (R92922123), 2005).

Backward Induction on the RT Tree (continued)

• For example, if K = 3, then a variance of

$$10.5436 \times 10^{-6}$$

will be added between the maximum and minimum variances at node (2,0) on p. 983.^a

• In general, the kth variance at node (i, j) is

$$h_{\min}^2(i,j) + k \frac{h_{\max}^2(i,j) - h_{\min}^2(i,j)}{K-1}, \quad k = 0, 1, \dots, K-1.$$

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

^aRepeated on p. 1003.



Backward Induction on the RT Tree (concluded)

- Suppose a variance falls between two of the K variances during backward induction.
- Linear interpolation of the option prices corresponding to the two bracketing variances will be used as the approximate option price.
- The above idea is reminiscent of the one in dealing with Asian options.^a

^aRecall p. 454.

Numerical Examples

- We next use the tree on p. 1003 to price a European call option with a strike price of 100 and expiring at date 3.
- Recall that the riskless interest rate is zero.
- Assume K = 2; hence there are no interpolated variances.
- The pricing tree is shown on p. 1006 with a call price of 0.66346.
 - The branching probabilities needed in backward induction can be found on p. 1007.





- Let us derive some of the numbers on p. 1006.
- A gray line means the updated variance falls strictly between h_{max}^2 and h_{min}^2 .
- The option price for a terminal node at date 3 equals $\max(S_3 100, 0)$, independent of the variance level.
- Now move on to nodes at date 2.
- The option price at node (2,3) depends on those at nodes (3,5), (3,3), and (3,1).
- It therefore equals

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

- Option prices for other nodes at date 2 can be computed similarly.
- For node (1, 1), the option price for both variances is
 0.1237 × 3.19054 + 0.7499 × 1.05240 + 0.1264 × 0.14573 = 1.20241.
- Node (1,0) is most interesting.
- We knew that a down move from it gives a variance of 0.000105609.
- This number falls between the minimum variance
 0.000105173 and the maximum variance 0.0001227 at
 node (2,-1) on p. 1003.

- The option price corresponding to the minimum variance is 0 (p. 1006).
- The option price corresponding to the maximum variance is 0.14573.
- The equation

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

is satisfied by x = 0.9751.

• So the option for the down state is approximated by

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

- The up move leads to the state with option price 1.05240.
- The middle move leads to the state with option price 0.48366.
- The option price at node (1,0) is finally calculated as
 0.4775 × 1.05240 + 0.0400 × 0.48366 + 0.4825 × 0.00362 = 0.52360.

- A variance following an interpolated variance may exceed the maximum variance or be lower than the minimum variance.
- When this happens, the option price corresponding to the maximum or minimum variance will be used during backward induction.^a
- This act tends to reduce the dynamic range of the variance, however.

^aCakici & Topyan (2000).

Numerical Examples (concluded)

- Worse, an interpolated variance may choose a branch that goes into a node that is *not* reached in forward induction.^a
- In this case, the algorithm fails.
- The RT algorithm does not have this problem.
 - This is because all interpolated variances are involved in the forward-induction phase.
- It may be hard to calculate the implied β_1 and β_2 from option prices.^b

^aLyuu & C. Wu (R90723065) (2005). ^bY. Chang (B89704039, R93922034) (2006).

Complexities of GARCH Models $^{\rm a}$

- The RT algorithm explodes exponentially even for moderate n.^b
- The mean-tracking tree of Lyuu and Wu (2005) guarantees explosion not to happen for n not too large.
 - That tree is similar to, but earlier than, the binomial-trinomial tree.^c
 - In fact, we can use the binomial-trinomial tree here, and everything goes through.^d

^aLyuu & C. Wu (**R90723065**) (2003, 2005).

^bRecall p. 979.

^cRecall pp. 764ff.

 $^{\rm d}{\rm Contributed}$ by Mr. Lu, Zheng-Liang (D00922011) on August 12, 2021.

Complexities of GARCH Models (continued)

• The next page summarizes the situations for many GARCH option pricing models other than NGARCH.

Complexities of GARCH Models (concluded)^a

Model	Explosion	Non-explosion
NGARCH	$\beta_1 + \beta_2 n > 1$	$\beta_1 + \beta_2(\sqrt{n} + \lambda + c)^2 \le 1$
LGARCH	$\beta_1 + \beta_2 n > 1$	$\beta_1 + \beta_2(\sqrt{n} + \lambda)^2 \le 1$
AGARCH	$\beta_1 + \beta_2 n > 1$	$\beta_1 + \beta_2(\sqrt{n} + \lambda)^2 \le 1$
GJR-GARCH	$\beta_1 + \beta_2 n > 1$	$\beta_1 + (\beta_2 + \beta_3)(\sqrt{n} + \lambda)^2 \le 1$
TS-GARCH	$\beta_1 + \beta_2 \sqrt{n} > 1$	$\beta_1 + \beta_2(\lambda + \sqrt{n}) \le 1$
TGARCH	$\beta_1 + \beta_2 \sqrt{n} > 1$	$\beta_1 + (\beta_2 + \beta_3)(\lambda + \sqrt{n}) \le 1$
Heston-Nandi	$\beta_1 + \beta_2 (c - \frac{1}{2})^2 > 1$	$\beta_1 + \beta_2 c^2 \le 1$
	$\& \ c \leq \frac{1}{2}$	
VGARCH	$\beta_1 + (\beta_2/4) > 1$	$\beta_1 \le 1$

^aY. C. Chen (**R95723051**) (2008); Y. C. Chen (**R95723051**), Lyuu, & Wen (**D94922003**) (2012).

Obtaining Profit and Loss of Delta Hedge

- Profit and loss of any hedging strategy should be calculated under the real-world probability measure.^a
- But hedging parameters such as delta should be computed under the risk-neutral measure.
- Say we want the distribution of profit and loss for the delta hedge under the GARCH model.
- If a tree is built for each sampled stock price to obtain the delta, the complexity will be astronomical.^b
- How to do it efficiently?^c

^aRecall p. 713. ^bAugustyniak, Badescu, & Guo (2021). ^cLu (D00922011), Lyuu, & Yang (D09922005) (2021).