# Fin 514: Financial Engineering II

## Lecture 16: Monte Carlo Methods for American options

Dr. Martin Widdicks

UIUC

Spring, 2018

# Outline

- In this lecture we will focus on adapting Monte Carlo methods to deal with early exercise features.
- We will focus upon an adaptation of the standard Monte Carlo method, described by Longstaff and Schwartz (2001)

# Overview on Monte Carlo methods

- One of the key unanswered questions in finance is how to value options with early exercise features by using a Monte Carlo method.
- Mathematically, the American option value $V_0$ is

$$V_0 = \max_{\tau} E_0^Q[e^{-r(\tau)} \max(S_\tau - K, 0)]$$

  where $\tau$ denotes a *stopping time*.
- The problem comes from the fact that Monte Carlo is a forward looking method. To use the sample paths we would have to test early exercise at each point in time on each sample path in order to determine what the optimal exercise strategy would be.
- This is incredibly time consuming and not practical, and simplistic approaches, such as the perfect foresight method where you simply choose the highest early exercise value during the lifetime of the option do not give acceptable approximations to the option value.

# Graphical explanation of perfect foresight

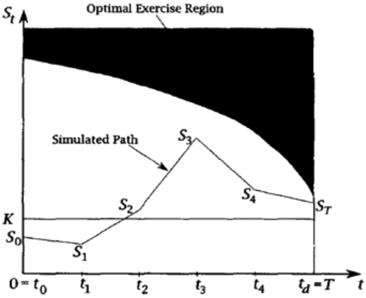


Fig. 1. Sample simulated path.

If you follow the optimal strategy (shaded) then you should only exercise at expiry, but you can get a better payoff from exercising at $t_3$. This is perfect foresight and does not value the option correctly

## Quiz

- Why is it important to be able to value options with early exercise features with Monte Carlo?

# Quiz

- Why is it important to be able to value options with early exercise features with Monte Carlo?
  Because in some cases we will have large numbers of underlying assets but American/call features and so we will have to use Monte Carlo

# Simple attempts

- Tilley (1993) made the first effort to adapt Monte Carlo methods to cope with early exercise features. The method involves a technique known as “bundling” where the paths are constructed as usual and, at each timestep, they are bracketed into regions of asset price.
- At expiry the option price is the average of the payoff from all the paths in the bundle. Working backwards through time, as the paths are known at each point it can be checked whether or not it was, on average, worth exercising in each bundle by comparing this to the discounted option price from the future time step.
- There are evident shortcomings with this approach:
  - First, all the paths have to be stored, which can cause computer memory problems.
  - More importantly the process overvalues the options.
  - Crucially, it is also very difficult to extend to options on multiple underlyings (usually the Monte Carlo method’s advantage over other numerical techniques).

# Early attempts II

- Realizing the main drawback in Tilley's method, Barraquand and Martineau (1995) adapted his approach so that the bundling was in terms of payoff value rather than underlying asset value. Payoff value only has one dimension and so extension to many underlyings does not create any undue problems.
- However, although not requiring as much memory as Tilley's approach, the approach still does not converge to the correct value and always underestimates the option value (see Boyle et al., 1997) and this estimation error can be serious.

# Broadie and Glasserman (1997)

- Broadie and Glasserman (1997) approach the Monte Carlo method as by creating upper and lower bounds. To create these bounds they use a “bushy tree effect ”to pursue sub- and super-optimal strategies. The super-optimal strategy is obtained by creating a tree whose possible next states are determined, by simulation, all the way to expiry.
- Then, in a similar vein to Tilley (1993) the option value at the previous time, $t$, is the maximum of the average of the values at $t + \Delta t$ discounted and the value from early exercise at $t$. This strategy does assume the investor has some foresight and so overvalues the option.

# Broadie and Glasserman (1997)

- The sub-optimal procedure entails using the $b$ possible paths at each time and, for each path,using the remaining $b-1$ paths to determine whether the option is continued or exercised. This exercise choice is then applied the initial path one was focusing on. All the possible combinations are then averaged at each timestep.
- This method is shown to be sub-optimal, for more details see Broadie and Glasserman (1997). These can then be combined to provide bounds for the put option value.

# Bushy tree

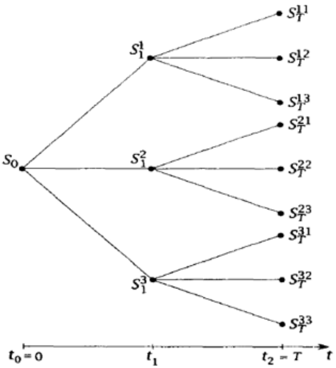


Fig. 2. Simulated tree for $b = 3$.

# Super-optimal technique

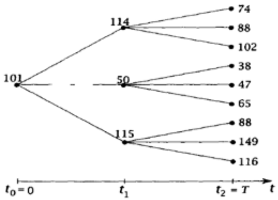


Fig. 3. Stock price tree.

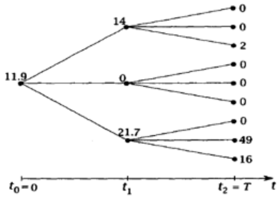


Fig. 4. $\Theta$ ('high') estimate.

# Sub-optimal technique

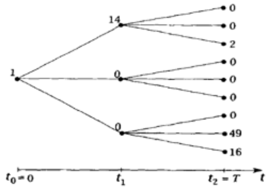


Fig. 5. Simple low estimate.

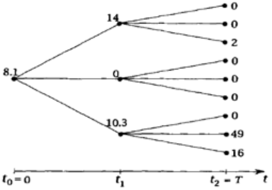


Fig. 6. $\theta$ ('low') estimate.

# Broadie and Glasserman (1997)

- As usual, the computational effort increases only linearly as more underlying assets are added. However, as the number of observation times is increased the calculations increase exponentially, I.e with n paths, d observation dates and b branches in the tree the effort is nbd (b here is typically quite large, e.g. 50)
- Thus, to estimate the value of a continuously, or even daily, observed option involves the use of extrapolation which is somewhat ad hoc if, as is the case here, the initial results are not converging at a known rate.

# More practical advances

- The most popular method for incorporating early exercise features in the Monte Carlo methods is by Longstaff and Schwartz (2001) which we will see in detail shortly. Its appeal is that it is simple to implement although there remain questions about its accuracy and efficiency.
- Another, more academically rigorous approach is the dual approach Haugh and Kogan (2004) which expresses the option pricing problem as a minimization problem, from which a tight upper bound on the price is calculated. Unfortunately, its calculation is often problematic, although there are practical approaches to circumvent this - see Andersen and Broadie (2004) for more details.

# Recap on problem

- The key calculations when valuing an American option are determining the early exercise value and the continuation value.
- The option value at any point in time will be the larger of these two values, e.g for a Put option:

$$V(S,t) = max(K - S, e^{-r\Delta t} E_t[V(S, t + \Delta t)])$$

- The early exercise value is simple to calculate at any point in time. The continuation value is typically determined by a backward iteration approach, where the value now is the discounted expected option value at the next instance in time.
- The key information we need to know is given the share price at time $t$, what is the expected option value at time $t + \Delta t$, $E_t[V(S, t + \Delta t)]$.
- This is simple to determine in binomial models but less so in Monte Carlo approaches.

# Longstaff and Schwartz (2001)

- The Longstaff and Schwartz method, essentially estimates the conditional expected option value at the next time step by simulating lots of paths but then simplifying the calculation of the expectation by carrying out a regression.
- The regression approximates the continuation value *on that path* as a function of the current value of the underlying asset.
- This gives an approximation for the continuation value that can then be compared to the early exercise value and then we know the option value at each point in time on each path.
- In terms of Monte Carlo pricing, all we actually need to know is the rule for early exercising, so we know when we receive the cash flows and the value of the option is the average of the discounted payoffs for each path.
- We will explain the method via an example and then describe the general method.

# Example

- We will attempt to value a Bermudan put option where exercise is possible now and at three future dates. $S_0 = 1, K = 1.1, r = 0.06$.
- The first step is to simulate some paths, the table below denotes the results for $N = 8$ and $M = 4$.

Stock price paths

| Path | $t = 0$ | $t = 1$ | $t = 2$ | $t = 3$ |
|---|---|---|---|---|
| 1 | 1.00 | 1.09 | 1.08 | 1.34 |
| 2 | 1.00 | 1.16 | 1.26 | 1.54 |
| 3 | 1.00 | 1.22 | 1.07 | 1.03 |
| 4 | 1.00 | .93 | .97 | .92 |
| 5 | 1.00 | 1.11 | 1.56 | 1.52 |
| 6 | 1.00 | .76 | .77 | .90 |
| 7 | 1.00 | .92 | .84 | 1.01 |
| 8 | 1.00 | .88 | 1.22 | 1.34 |

- We need to use this information to determine the continuation value at each point in time ($e^{-r\Delta t} E_t[V(S, t + \Delta t)]$) for each path. To do this we will construct a Cash Flow Matrix at each point in time.

# Continuation value at t=2

- The table below denotes the cash flows at $t = 3$ assuming that we held the option that far:

Cash-flow matrix at time 3

| Path | $t = 1$ | $t = 2$ | $t = 3$ |
|---|---|---|---|
| 1 | — | — | .00 |
| 2 | — | — | .00 |
| 3 | — | — | .07 |
| 4 | — | — | .18 |
| 5 | — | — | .00 |
| 6 | — | — | .20 |
| 7 | — | — | .09 |
| 8 | — | — | .00 |

- The next step is to attempt to find a function that describes the continuation value at time 2 as a function of the value of $S$ at time 2. We think of the continuation value as $e^{-r\Delta t}E_t[V_{t+\Delta t}]$.
- To do this we use a regression technique, that takes the stock prices at time 2 as the "$X$" values and the discounted payoffs (or continuation values) at time 3 as the "$Y$" values.

# Continuation value at t=2

Regression at time 2

| Path | $Y$ | $X$ |
|---|---|---|
| 1 | $.00 \times .94176$ | 1.08 |
| 2 | — | — |
| 3 | $.07 \times .94176$ | 1.07 |
| 4 | $.18 \times .94176$ | .97 |
| 5 | — | — |
| 6 | $.20 \times .94176$ | .77 |
| 7 | $.09 \times .94176$ | .84 |
| 8 | — | — |

- Note that the regression is only carried out on paths that are in the money at time 2.
- The regression here is simple where $Y$ is regressed upon $X$ and $X^2$ (the actual scheme is slightly more sophisticated). In this particular example: $Y = -1.070 + 2.938X - 1.813X^2$, and so we can use this to estimate the continuation value for each of the current share prices ($X$ in the regression).
- For example for path 1, where $S = 1.08$, the regression formula gives $Y$ (the continuation value) to be 0.0369.

# Quiz

Why don't we regress with out-of-the money paths?

- It makes the regression wrong
- The only interesting paths are the in the money ones so this saves time.
- Some other explanation

# Quiz

Why don't we regress with out-of-the money paths?

- It makes the regression wrong
- The only interesting paths are the in the money ones so this saves time. CORRECT
- Some other explanation

# Quiz

Do we want the regression to precisely predict $V(S, t + \delta t)$ on that path?

- Yes
- No.
- Maybe

# Quiz

Do we want the regression to precisely predict $V(S, t + \delta t)$ on that path?

- Yes
- No. CORRECT
- Maybe

# Option value at t=2

Optimal early exercise decision at time 2

| Path | Exercise | Continuation |
|---|---|---|
| 1 | .02 | .0369 |
| 2 | — | — |
| 3 | .03 | .0461 |
| 4 | .13 | .1176 |
| 5 | — | — |
| 6 | .33 | .1520 |
| 7 | .26 | .1565 |
| 8 | — | — |

- This then allows you to decide at which points in time you would exercise and thus determine the cash flows at $t = 2$ (below). Notice that for each path, if you exercise at $t = 2$ then you do not also exercise at $t = 3$.

Cash-flow matrix at time 2

| Path | $t = 1$ | $t = 2$ | $t = 3$ |
|---|---|---|---|
| 1 | — | .00 | .00 |
| 2 | — | .00 | .00 |
| 3 | — | .00 | .07 |
| 4 | — | .13 | .00 |
| 5 | — | .00 | .00 |
| 6 | — | .33 | .00 |
| 7 | — | .26 | .00 |
| 8 | — | .00 | .00 |

# Quiz

- Why are cash flows at subsequent times set to zero if you exercise at $t = 2$?

# Quiz

- Why are cash flows at subsequent times set to zero if you exercise at $t = 2$? Because you have exercised at t = 2 and so receive no other cash flows.

# Continuation value at t=1

- We can apply the same process to $t = 1$, for each of the paths that are in the money we regress the discounted future cash flows ($Y$) on the current value of the underlying asset ($X$), where $X$ and $Y$ are as given below:

Regression at time 1

| Path | $Y$ | $X$ |
|---|---|---|
| 1 | $.00 \times .94176$ | 1.09 |
| 2 | — | — |
| 3 | — | — |
| 4 | $.13 \times .94176$ | .93 |
| 5 | — | — |
| 6 | $.33 \times .94176$ | .76 |
| 7 | $.26 \times .94176$ | .92 |
| 8 | $.00 \times .94176$ | .88 |

- The regression equation here is $Y = 2.038 - 3.335X + 1.356X^2$ and again we use this to estimate the continuation value and decide on an early exercise strategy.
- The next table compares the two values and the final table denotes the early exercise or stopping rule.

# Stopping rule

Optimal early exercise decision at time 1

| Path | Exercise | Continuation |
|---|---|---|
| 1 | .01 | .0139 |
| 2 | — | — |
| 3 | — | — |
| 4 | .17 | .1092 |
| 5 | — | — |
| 6 | .34 | .2866 |
| 7 | .18 | .1175 |
| 8 | .22 | .1533 |

- The early exercise strategy is as follows:

Stopping rule

| Path | $t = 1$ | $t = 2$ | $t = 3$ |
|---|---|---|---|
| 1 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0 |
| 3 | 0 | 0 | 1 |
| 4 | 1 | 0 | 0 |
| 5 | 0 | 0 | 0 |
| 6 | 1 | 0 | 0 |
| 7 | 1 | 0 | 0 |
| 8 | 1 | 0 | 0 |

- From this we can then value the option, by forming the final cash flow matrix from this rule.

# Option value

Option cash flow matrix

| Path | $t = 1$ | $t = 2$ | $t = 3$ |
|---|---|---|---|
| 1 | .00 | .00 | .00 |
| 2 | .00 | .00 | .00 |
| 3 | .00 | .00 | .07 |
| 4 | .17 | .00 | .00 |
| 5 | .00 | .00 | .00 |
| 6 | .34 | .00 | .00 |
| 7 | .18 | .00 | .00 |
| 8 | .22 | .00 | .00 |

- So the option value is the average of the discounted cash flows, so in this case:

$$\begin{aligned} V_0 &= \frac{1}{8}\left(0 + 0 + 0.07e^{-3r} + 0.17e^{-r} + 0 + 0.34e^{-r}\right. \\ &\quad \left. + 0.18e^{-r} + 0.22e^{-r}\right) \\ &= 0.1144 \end{aligned}$$

# More sophisticated regression

- In general the regression here $Y = a_1 + a_2 X + a_3 X^2$ is not going to be satisfactory, especially as we will have far more than 8 paths when attempting to find the functional form of the continuation values.
- In fact the general form of $Y$ is:

$$Y = \sum_{j=0}^{M} a_j F_j(X)$$

- The user decides upon $M$ the number and the functional form $F_j(X)$ (in our example $F_j(X) = X^j$) where $F_j(X)$ is called a basis function.
- The suggestion of Longstaff and Schwartz is to choose basis functions described by Laguerre polynomials to provide the best fit. However, you are free to choice whichever basis functions you want (e.g. Polynomial as in the example, Chebyshev, Hermite etc.)

# More sophisticated regression

- The Laguerre polynomials are given by:

$$\begin{aligned}
F_0(X) &= \exp(-X/2) \\
F_1(X) &= \exp(-X/2)(1-X) \\
F_n(X) &= \exp(-X/2)\frac{e^X}{n}\frac{d^n}{dX^n}(X^n e^{-X})
\end{aligned}$$

- Then the least squares approach approximates the constants $a_j$, and when we have these values we can use them to predict the continuation value for each value of $S$ at every point in time.

# General procedure: slightly different from LS

1. Decide on the number of sample paths $N$, the number of basis functions for the regression, $M$, and the type of basis functions $F_j(X)$ and the number of observation dates $d$.
2. Draw $Nd$ Normally distributed random numbers and simulate the sample paths for the underlying asset at each point in time $S^n_{t_i}, \qquad 1 \leq i \leq d, \qquad 1 \leq n \leq N$
3. At expiry $t = t_d$, record the option values $V^n(t_d)$ which for a put are $\max(K - S^n_{t_d}, 0)$.
4. Move back to $t = t_{d-1}$ for each path where $S^n_{t_{d-1}} < K$ calculate the continuation value as $CV^n(t_{d-1}) = e^{-r(t_d - t_{d-1})}V^n(t_d)$. From these values perform the regression to determine the functional form of the continuation value, $Y(S^n)$ where $S^n$ is the value of the underlying asset on path $n$.
5. Recalculate the continuation value as $CV^n(t_{d-1}) = Y(S^n_{t_{d-1}})$

# General procedure

6. For every path $(n)$ calculate the option value where if the continuation value $CV^n(t_{d-1}) < K - S^n_{t_{d-1}}$ then $V^n(t_{d-1}) = K - S^n_{t_{d-1}}$, otherwise $V^n(t_{d-1}) = e^{-r(t_d - t_{d-1})} V^n(t_d)$.
7. Repeat this process for the previous time step until you have $V^n(t_0)$ for all $n$. Note that in general to calculate $CV^n(t_i)$ (or $Y$) before the regression in this algorithm:

$$CV^n(t_i) = e^{-r(t_{i+1} - t_i)} V^n(t_{i+1})$$

8. The option value $V_0$ is then

$$V_0 = \frac{1}{N} \sum_{i=1}^{N} V^i(t_0)$$

## Quiz

- Why do steps 7 and 8 correctly describe the option value?

# Quiz

- Why do steps 7 and 8 correctly describe the option value? What you end up with is the payoff from the option, and by the fundamental theorem the current value of the option is its payoff discounted at the risk-free rate. Then these option values are averaged and by the central limit theorem this will converge to the correct option value.

# More sources of uncertainty

- When you have more than one underlying asset in order to perform the regression you need to have basis functions in all of the underlying assets as well as in the cross terms between them (i.e. in $S_1$, $S_2$ and $S_1 S_2$).
- This means that the number of basis functions will increase exponentially as you increase the number of underlying assets, although it is not necessary in practice to have too large a number of basis functions.

# How well does it perform?

- Longstaff and Schwartz provide proofs that as $M \to \infty$ and $N \to \infty$ the option value obtained from their scheme converges to the theoretical value.
- This isn't much use for practical considerations as you will be limited by how many basis functions you can calculate and how many simulations you can perform.
- There have been a few investigations into the method and the view of the methods performance are mixed: Broadie and Detemple (2004) say regression methods 'often incur unknown approximation errors and are limited by a lack of error bounds'.

# How well does it perform?

- A detailed appraisal of this technique by Moreno and Navas (2003) investigates the use of various polynomial fits and numbers of basis functions.
- It is not clear that increasing the number of basis functions actually increases the accuracy of the method and there is no real difference from using different types of basis functions (e.g. Chebyshev rather than Laguerre)
- For more complicated derivative pricing problems the trend is even less clear, sometimes errors can increase as you add more basis functions (too many essentially fits the stochastic values of $S$ exactly)
- In general, the method will provide good estimates but will be difficult to assess exactly how accurate it is.

# Conclusion

- We have looked at binomial and Monte Carlo methods for valuing opions and where their strengths and weaknesses lie.
- Monte Carlo methods are better at dealing with more sources of uncertainty but cannot be easily adapted to consider American option problems.
- We have then introduced a method for valuing options with early exercise features using simulation.
- The main idea is to estimate the continuation value (as a function of the current underlying asset price) by performing a least squares regression.
- The method converges to the correct option price but research shows that it is unclear how well the method performs and how to estimate the error when using a finite number of paths/ basis functions.