BI Norwegian Business School - campus Oslo

GRA 19502

Master Thesis

Component of continuous assessment: Forprosjekt, Thesis MSc

Pricing American-Style Options by Monte Carlo Simulation

Navn: Bjarte Joleik, Vilde Rivers Marhaug

Start: 01.01.2018 09.00

Finish: 15.01.2018 12.00

Abstract

Within the past decades, several Monte Carlo simulation-based approaches have been proposed to address the problem of pricing American-style derivatives. In more recent years, least squares regression-based Monte-Carlo methods have been proposed specifically for the computation of American option prices. The purpose of our thesis is to replicate and confirm Tompaidis and Yang's (2014) evaluation of Longstaff and Schwartz's (2001) Least Squares Monte Carlo (LSM) algorithm by comparing its performance of Ordinary Least Squares against other numerical methods. This preliminary report will lay the foundation for our thesis and includes an introduction to the topic, a literature review, theory related to American simulation based option pricing, a thesis progression plan, and concluding remarks. Since the LSM method of Longstaff and Schwartz (2001) is the fundamental building block in the work of Tompaidis and Yang (2014), we focus on theory related to LSM. We also include a MATLAB-function of the LSM algorithm that will work as a framework for our replication of the methods presented by Tompaidis and Yang (2014).

Content

1. Introduction

One of the most important problems in option pricing theory is the valuation and optimal exercise of American-style options. These early-exercise derivatives can be exercised not only at their point of maturity, but at any moment preceding. Their valuation typically involves solving an optimal stopping problem. For simple vanilla options where only one factor affects the value of the option, the optimal stopping problem can be solved efficiently using conventional numerical procedures. However, when more than one factor affects the value of the option, pricing procedures become exceptionally complicated.

1.1 Problem Description

The Black and Scholes formula is commonly known as a closed-form solution for valuing European options. In contrast, no closed-form solution exists for valuing American options. When only one factor affects the value of an American option, it is conventionally valued by lattice methods such as binomial- and trinomial trees as well as finite difference methods. However, these methods become difficult to evaluate accurately when multiple stochastic factors affect the value of the option. When problems with multi-dimensional features are considered, Monte Carlo methods often give better results since the convergence rate in Monte Carlo simulations is independent of the number of stochastic state variables.

The major drawback of Monte Carlo simulation is its difficulty of dealing with the early exercise feature embedded in American options. The problem of using simulation based methods to price American options results from the difficulty in applying a forward based procedure to a problem that requires a backward procedure to be solved. Because of the early exercise feature embedded in American options, we need to know the value of the option at intermediate times between the start of the simulation and when the option expires. With Monte Carlo, this information is hard to obtain. Therefore, even though Monte Carlo is capable of handling multi-factor problems, once we need to solve a problem backwards, it becomes difficult to implement. Several researchers have provided ways of valuing American options when Monte Carlo simulation is used. In 2001, Longstaff and Schwartz proposed the Least Squares Monte Carlo (LSM) method. Their approach involves using a least-squares regression analysis to determine the best-fit relationship between the continuing value and the values of relevant variables at each time an early exercise decision must be made. The method has achieved much popularity because of its intuitive regression-based approach to pricing American option. In recent years, Tompaidis and Yang (2014) have evaluated the performance of LSM against quantile regression, Tikhonov Regularization, Matching Project Pursuit (MPP), and Classification and Regression Trees (CART). They find that LSM is inclined to over fit in several instances such as when the frequency of exercise increases or when a low number of simulation paths is used. Additionally, their analysis find that several of the other methods outperform LSM when European option prices is included in the polynomial basis functions.

1.2 Research Question

The objective of this thesis is to replicate (in some parts) and confirm the analysis of Tompaidis and Yang (2014) by comparing its performance to other numerical methods, and to extend and improve the methodology along different directions by finding other basis functions that can improve the pricing accuracy. The research question for our thesis is: What are strengths and weaknesses of methods proposed by Tompaidis and Yang (2014), how do they differ in terms of efficiency, robustness and precision, and what are improvements to be done.

1.3 Research Method

The performance of Ordinary Least Squares (OLS) in the Least Squares Monte Carlo (LSM) algorithm will be compared to several alternative methods such as Tikhonov Regularization, Matching Project Pursuit (MPP), and Classification and Regression Trees (CART). A set of five test cases that were introduced by Fu, Laprise, Madan, Su, and Wu (2001) will be used as a benchmark. The test cases include plain vanilla options with and without continuous dividends, jump-diffusion option, Parisian option, lookback option, and moving window Asian option. It should be studied whether each option’s obtained price converges to its true price.

1.4 Preliminary Report Structure

The rest of the preliminary report is organized as follows. Chapter 2 contains a literature review of studies on the topic of simulation based option pricing. Chapter 3 outlines theory related to American simulation based option pricing. Chapter 4 presents the MATLAB function of the LSM algorithm that will work as a framework for our replication of the pricing functions. Chapter 5 outlines a plan for our progression of the thesis. Chapter 6 concludes the preliminary report.

2. Literature Review

The modern version of the Markov Chain Monte Carlo method was invented in the late 1940s by Stanislaw Ulam, and Phelim Boyle was among the first to introduce Monte Carlo simulation into finance by it proposing it for the study of European option prices in 1977.

In the 1990's, the first approaches were presented to which Monte Carlo simulation can be used to value American-style options. Tilley (1993) was the first who attempted to use Monte Carlo simulation to value American options by using a bundling technique and a backward induction algorithm to determine the early exercise boundary. With improvements on the basic idea of Tiller, Carriere (1996) presents a backward induction algorithm and applies it to calculate an early exercise premium. He shows that the estimation of the early exercise decision rule should be equivalent to the estimation of a series of conditional expectations. In his algorithm, the conditional expectations are estimated using nonparametric least squares regression of spline functions. Other early work includes Grant, Vora and Weeks (1997) and Broadie and Glasserman (1997) who considers more general path-dependent options such as Asian options. Broadie and Glasserman shows how to price Asian options by Monte Carlo, but their method does not focus on an optimal exercise strategy. Instead, they compute a confidence interval and generate two biased estimators; an upper (biased high) and a lower (biased low) bound that converges asymptotically (and unbiasedly) to the true price of an American option.

Carriere's idea was further developed by Tsitsiklis and Van Roy (1999), Tsitsiklis and Van Roy (2001), and Longstaff and Schwartz (2001) who uses least squares regression to approximate the continuation value function by its projection on the linear span of a set of functions. Tsitsiklis and Van Roy (2001) use all the simulated paths to estimate the continuation value. In contrast to Tsitsiklis and Van Roy (2001), Longstaff and Schwartz (2001) only use price paths that are in-the-money to increase the efficiency of the algorithm. In their method, they apply least squares regression in which the explanatory variables are certain polynomial functions and estimate the continuation values of several types of derivatives.

The convergence properties of the LSM algorithm have been studies by Clément, Lamberton and Protter (2002). They demonstrate that the estimated conditional expectation approaches the true conditional expectation as the number of basis functions goes to infinity. Glasserman and Yu (2004) study the convergence rate of the algorithm when the number of basis functions and the number of paths increase simultaneously. They show that in certain cases, to guarantee that the option price converges to its true price, the number of paths must grow exponentially with the number of polynomial basis functions when the underlying state variable follows Brownian motion. If the underlying variable follows geometric Brownian motion, the number of paths must grow faster than exponential to guarantee convergence.

Other studies of the LSM algorithm include the following authors. Moreno and Navas (2003) analyze the robustness of the algorithm with respect to basis function selection. Gamba (2003) extends it to value real options. Rasmussen (2005) and Fouque and Han (2007) attempts to improve the efficiency of LSM by including control variates.

In recent years, Tompaidis and Yang (2014) critically evaluate the performance of LSM against quantile regression, Tikhonov regularization, Matching Projection Pursuit (MPP), a modified version of MPP, and Classification and Regression Trees (CART. They find that LSM is inclined to over fit in several instances such as when the frequency of exercise increases or when a low number of simulation paths is used. Additionally, their analysis find that several of the other methods outperform LSM when European option prices is included in the polynomial basis functions.

3. American Simulation Based Option Pricing Theory

The first part of chapter 3 outlines the theory behind the LSM algorithm by Longstaff and Schwartz (2001). The second part describes a variance reduction technique that can be implemented to reduce computational time in the pricing functions.

3.1 Least Squared Monte Carlo (LSM) Algorithm

Tompaidis and Yang (2014) suggest using the following notation and approach for the LSM algorithm: $S_t^{(i)}$ is the value of the state variables at time *t* along path *i*; *h* is the option payoff; *V* is the option value; $\{t_i\}_{i=0}^{N}$ are the possible exercise times.

3.1.1 Price Path Simulation

The LSM algorithm starts by simulating *M* possible price paths a stock may follow during a specified time span *t*. The price is logarithmic and follows a geometric Brownian motion stochastic process. It can be computed by the following equation.

$$S(t_{j+1}) = S(t_j)\, exp\left[\left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \sigma\varepsilon\sqrt{\Delta t}\right]$$

In the equation, $S(t_j)$ is stock price at time j; μ is the expected return in a risk-neutral world; σ is the volatility; ε is a random number drawn from the standard normal distribution; Δt is the length of time interval. Below is an illustration of a simulation.

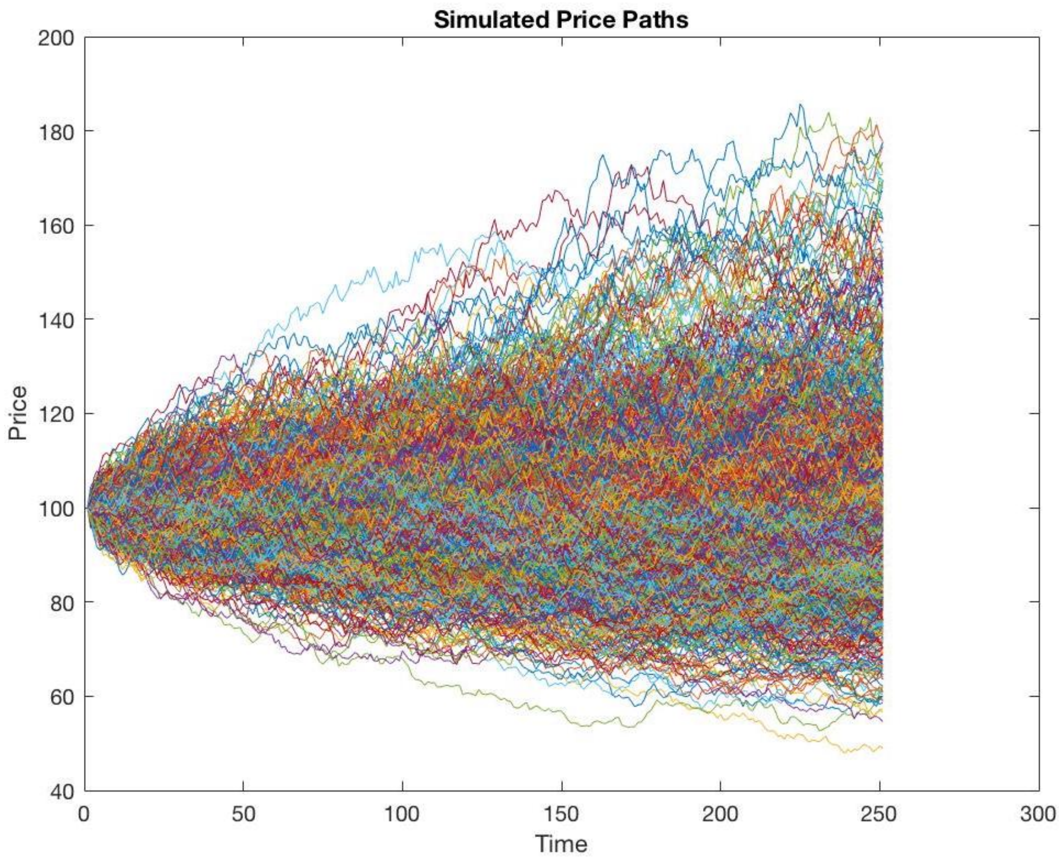

Figure 1: Simulation of Price Paths (S=100, K=90, r=0.05, sigma=0.2, D=0, T=1, NSteps=250, NSims=1000)

3.1.2 Payoff Computation

For all *M* paths and all exercise time points *t,* the options cash flow matrix *V* is computed. At the terminal value t_N, set the option value *V* equal to the payoff.

$$V\left(S_{t_N}^{(m)}, t_N\right) = h\left(S_{t_N}^{(m)}, t_N\right), m = 1, \dots, M$$

3.1.2 Conditional Expectation Value Computation

The algorithm then proceeds backwards by projecting the expectation of the subsequent discounted cash flows onto the basis functions for the paths where the option is in the money at time t_{j-1}.

For the set of paths $\{i_l\}_{l=1}^{L}$, for which the option is in-the-money; i.e., $h\left(S_{t_{N-1}}^{i_l}, t_{N-1}\right) > 0$, find coefficients $a_j^*(t_{N-1})$ to minimize the norm

$$\left\| \sum_{j=1}^{N_b} a(t_{n-1}) \begin{matrix} \varphi_j\left(S_{t_{N-1}}^{(i_1)}\right) \\ \varphi_j\left(S_{t_{N-1}}^{(i_2)}\right) \\ \vdots \\ \varphi_j\left(S_{t_{N-1}}^{(i_L)}\right) \end{matrix} - e^{-r(t_N - t_{N-1})} \begin{pmatrix} V\left(S_{t_N}^{(i_1)}, t_N\right) \\ V\left(S_{t_N}^{(i_2)}, t_N\right) \\ \vdots \\ V\left(S_{t_N}^{(i_L)}, t_N\right) \end{pmatrix} \right\|$$

Here, ‖·‖ is the vector norm, and $\left({\varphi_j}_{t_{j=1}}^{(N_b)}\right)$ is the basis function. The fitted values are chosen as the expected continuation values. Longstaff and Schwartz (2001) estimate the conditional expectation value directly by least squares regression, while Tompaidis and Yang (2014) use alternative methods.

3.1.4 Optimal Exercise Decision

To make the optimal exercise decision, we compare the vector norm ‖·‖ with the value for immediate exercise $V\left(S_{t_{j-1}}^{i}\right)$ at each path.

$$V\left(S_{t_{N-1}}^{(m)}, t_{N-1}\right) = \begin{cases} h\left(S_{t_{N-1}}^{(m)}, t_{N-1}\right) & if\ h\left(S_{t_{N-1}}^{(m)}, t_{N-1}\right) \geq \sum_{j=1}^{N_b} a_j^*(t_{N-1})\, \varphi_j\left(S_{t_{N-1}}^{(m)}\right) \\ e^{-r(t_N - t_{n-1})} V\left(S_{t_N}^{(m)}, t_N\right) & otherwise \end{cases}$$

3.1.5 Backwards Recursion

We recursively use this algorithm for possible exercise times $t_{N-2}, t_{N-3}, \ldots$, until time t_0. That is the option price.

3.2 Variance Reduction Techniques

If the stochastic processes for the variables underlying a derivative are simulated as by Monte Carlo, a very large number of trials is usually necessary to estimate the value of the derivative with reasonable accuracy. This is very expensive in terms of computation time (Hull, 2015). This chapter presents different variance reduction techniques that can lead to savings in computational time in the pricing functions.

3.2.1 Antithetic Variable Technique

In a Monte Carlo simulation, the set of possible outcomes are determined by a random draw from the normal distribution. Because of this random draw, it often becomes the case that after many replications, the set of outcomes does not necessarily contain a complete range of all the possible outcomes (Brooks, 2014). The antithetic variable technique is used as an instrument in simulation to reduce the variance. A simulation trial involves calculating two values; the first is calculated as usual and the second is calculated by adding a negative sign in front of all the numbers drawn from the normal distribution. By doing this, we get complementary values for all the original values. Thus, the total range of possible outcomes is better represented.

4. MATLAB function

In this chapter, we include a MATLAB-function that will work as a framework for our replication of the methods presented by Tompaidis and Yang (2014). We will continue to develop our own MATLAB-code based on the function presented, relevant theory, and existing statistical packages made for MATLAB.

4.1 LSM for an American Call Option with Continuous Dividends

This function prices an American vanilla call option with continuous dividends using LSM. The function is originally coded by Mark Hoyle (2016), although we have changed its syntax and modified it to price an American call option with continuous dividends. We have also implemented the Antithetic Variable Technique.

```
function LSMAmCallContDiv(S0,K,D,r,sigma,T,NSteps,NSims)
%% LSM FOR AN AMERICAN CALL OPTION WITH CONTINOUS DIVIDENDS
% This function prices an American vanilla call option with
% continous dividends using LSM. The function works
% recursively by using simulated price paths to determine
% the conditional expected continuation value. The optimal
% stopping strategy is defined as when the intrinsic value
% of the options is greater than the conditional expected
% value of continuation.

% SYNTAX
% S0: Initial asset price
% K: Strike price
% D: Continuous dividend yield
% r: Risk-free rate
% sigma: Volatility
% T: Time to maturity (years)
% NSteps: Number of time steps
% NSims: Number of simulations

% OUTPUT
% Price
% Standard error

%% Generating asset paths
diff = randn(NSims,NSteps);
SPaths = zeros(NSims, 1+NSteps);
SPaths(:,1) = S0; % Stock price
dt = T/NSteps; % Length of time interval
nudt = (r-D-0.5*sigma^2)*dt;
sidt = sigma*sqrt(dt);

for i=1:NSims
    for j=1:NSteps
    SPaths(i,j+1)=SPaths(i,j)*exp(nudt + sidt*diff(i,j));
    end
end

%% Generating cash flow matrix
CF=zeros(NSims,NSteps+1);
CF(:,end)=max((SPaths(:,end)-K),0); % Terminal values

% Computing cash flow matrix
for j=NSteps:-1:1 % Since exercise at t=0 possible

        SpITM=find(SPaths(:,j)>K); % Locating paths that are
in the money
        X=SPaths(SpITM,j); X1=X/S0;
        Y=CF(SpITM,j+1)*exp(-r*dt); % Discounting cash flow
        LP=[ ones(size(X1)) (1-X1) 1/2*(2-4*X1-X1.^2)]; %
Weighted Laguerre polynomials
        Reg=pinv(LP)*Y; % Linear regression
        ExpContV=LP*Reg; % Conditional expected continuation
value
        ImExV=X-K; % Immediate exercise Value
        IdEx=find(ImExV > ExpContV); % Identify immediate
exercise
        Cont=setdiff(SpITM,SpITM(IdEx));
        CF(SpITM(IdEx),j)=max(X(IdEx)-K,0);
```

```
        CF(SpITM(IdEx),j+1:end)=0; % Generates zeros
        CF(Cont,j)=exp(-r*dt)*CF(Cont,j+1);
end

%% Generate dicount rate matrix
DisM=zeros(NSims,NSteps+1);
for i=1:NSims
    for j=2:NSteps+1
        DisM(i,j)=exp(-r*(j-1)*dt);
    end
end

%% Computing price per path
PVCF=CF.*DisM; % Present value of all cash flows
P=max(PVCF,[],2); % Prices of each run

%% Antithetic Variable Technique
diffa=-1*(diff);
SPaths2=zeros(NSims,NSteps+1); % Initialize matrix
SPaths2(:,1)=S0; % Each trial starts at S0
for i=1:NSims
    for j=1:NSteps
SPaths2(i,j+1)=SPaths2(i,j)*exp(nudt+sigma*sqrt(dt)*(diffa(i
,j))); %Creates a set of anthetic price paths
    end
end

CFa=zeros(NSims,NSteps+1); % Creating antithetic cash flow
matrix
CFa(:,end)=max((SPaths2(:,end)-K),0); % Computing end values

% Generating antithetic cash flow matrix
for j=NSteps:.1:1
        SpITMa=find(SPaths2(:,j)>K);
        Xa=SPaths2(SpITMa,j); X2=Xa/S0;
        Ya=CFa(SpITMa,j+1)*exp(-r*dt);
        LPa=[ ones(size(X2)) (1-X2) 1/2*(2-4*X2-X2.^2)];
        Rega=pinv(LPa)*Ya;
        ExpContVa=CMa*Rega;
        IntVa=Xa-K;
        IdExa=find(IntVa > ExpContVa);
        Conta=setdiff(SpITMa,SpITMa(IdExa));
        CFa(SpITMa(IdExa),j)=max(Xa(IdExa)-K,0);
        CFa(SpITMa(IdExa),j+1:end)=0;
        CFa(Conta,j)=zeros(length(Conta),1);
end

PVCFa=CFa.*DisM;
Pa=max(PVCFa,[],2); %Prices of each run

%% Computing option price and standard error
AllP=cat(1,P,Pa); % Creates one vector of prices and
anthetic prices
Price=(sum(AllP)/(NSims*2)); % Weighted average of vector
containing all prices
StdErr=std(AllP)/sqrt(NSims*2);
display (Price); % Display price
display (StdErr); % Display standard error
end
```

5. Thesis Progression Plan

January	- Improve function presented in the preliminary report.
February	- Program other LSM functions to price options in test cases. - Research, program, and test other potential pricing models to be included in the thesis.
March	- Compare performance of OLS to Tikhonov Regularization, MPP and CART. - Write, edit and review thesis drafts. - Review MATLAB functions.
April-June	- Write, edit, and review first draft of the thesis.
July	- Deliver draft to thesis supervisor for comments and suggestions.
July-August	- Implement suggestions from supervisor.
September	- Hand in thesis.

6. Conclusion

Through this report we have briefly described the research objectives for our thesis. Particularly, we will attempt to replicate and confirm Tompaidis and Yang's (2014) evaluation of Longstaff and Schwartz's (2001) LSM algorithm by comparing its performance of OLS regression to several other numerical methods. We expect that we will reach conclusions that will allow us to evaluate the strengths and weaknesses of the methods, and to assess their efficiency, robustness, and precision. We will also attempt to improve the current methodology of Tompaidis and Yang (2014) by introducing other numerical procedures.

7. Bibliography

Barraquand, J., & Martineau, D. (1995). Numerical valuation of high dimensional multivariate American securities. *Journal of financial and quantitative analysis, 30*(3), 383-405.

Boyle, P., Broadie, M., & Glasserman, P. (1997). Monte Carlo methods for security pricing. *Journal of economic dynamics and control, 21*(8), 1267-1321.

Boyle, P. P. (1977). Options: A monte carlo approach. *Journal of financial Economics, 4*(3), 323-338.

Broadie, M., & Glasserman, P. (1997). Pricing American-style securities using simulation. *Journal of economic dynamics and control, 21*(8), 1323-1352.

Broadie, M., & Kaya, O. (2004). *Exact simulation of option greeks under stochastic volatility and jump diffusion models.* Paper presented at the Simulation Conference, 2004. Proceedings of the 2004 Winter.

Brooks, C. (2014). *Introductory econometrics for finance*: Cambridge university press.

Carriere, J. F. (1996). Valuation of the early-exercise price for options using simulations and nonparametric regression. *Insurance: mathematics and Economics, 19*(1), 19-30.

Clément, E., Lamberton, D., & Protter, P. (2002). An analysis of a least squares regression method for American option pricing. *Finance and Stochastics, 6*(4), 449-471.

Cox, J. C., Ross, S. A., & Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of financial Economics, 7*(3), 229-263.

Fouque, J.-P., & Han, C.-H. (2007). A martingale control variate method for option pricing with stochastic volatility. *ESAIM: Probability and Statistics, 11*, 40-54.

Fu, M. C., Laprise, S. B., Madan, D. B., Su, Y., & Wu, R. (2001). Pricing American options: A comparison of Monte Carlo simulation approaches. *Journal of Computational Finance, 4*(3), 39-88.

Gamba, A. (2003). Real options valuation: A Monte Carlo approach.

Gamba, A. (2003). Real options valuation: A Monte Carlo approach.

Glasserman, P. (2013). *Monte Carlo methods in financial engineering* (Vol. 53): Springer Science & Business Media.

Glasserman, P., & Yu, B. (2004). Number of paths versus number of basis functions in American option pricing. *The Annals of Applied Probability, 14*(4), 2090-2119.

Grant, D., Vora, G., & Weeks, D. (1997). Path-dependent options: Extending the Monte Carlo simulation approach. *Management Science, 43*(11), 1589-1602.

Hull, J. C., & Basu, S. (2016). *Options, futures, and other derivatives*: Pearson Education India.

Longstaff, F. A., & Schwartz, E. S. (2001). Valuing American options by simulation: a simple least-squares approach. *The review of financial studies, 14*(1), 113-147.

Metropolis, N., & Ulam, S. (1949). The monte carlo method. *Journal of the American statistical association, 44*(247), 335-341.

Moreno, M., & Navas, J. F. (2003). On the robustness of least-squares Monte Carlo (LSM) for pricing American derivatives. *Review of Derivatives Research, 6*(2), 107-128.

Moreno, M., & Navas, J. F. (2003). On the robustness of least-squares Monte Carlo (LSM) for pricing American derivatives. *Review of Derivatives Research, 6*(2), 107-128.

Rasmussen, N. S. (2005). Control variates for Monte Carlo valuation of American options.

Tilley, J. A. (1999). Valuing American options in a path simulation model.

Tompaidis, S., & Yang, C. (2014). Pricing American-style options by Monte Carlo simulation: alternatives to ordinary least squares.

Tsitsiklis, J. N., & Van Roy, B. (2001). Regression methods for pricing complex American-style options. *IEEE Transactions on Neural Networks, 12*(4), 694-703.

Hoyle, M. (September 1, 2016). Pricing American Options. Retrieved January 1, 2018, from https://se.mathworks.com/matlabcentral/fileexchange/16476-pricing-american-options?focused=6781443&tab=function&s_tid=gn_loc_drop