Asian Journal of Economics, Business and Accounting

21(11): 34-39, 2021; Article no.AJEBA.70962
ISSN: 2456-639X

Using Logistic Regression of Machine Learning Method to Evaluate American Options

Yung Hsin Lee[1*]

[1]*Department of International Business, Tamkang University No.151,Yingzhuan Rd., Danshui Dist., New Taipei City 25137,Taiwan (R.O.C).*

Author's contribution

The sole author designed, analysed, interpreted and prepared the manuscript.

Article Information

DOI: 10.9734/AJEBA/2021/v21i1130439
Editor(s):
(1) Dr. Ivan Markovic, University of Nis, Serbia.
Reviewers:
(1) Haoran Wei, University of Texas, USA.
(2) Shruti Mohapatra, Sri Sri University, India.
Complete Peer review History: https://www.sdiarticle4.com/review-history/70962

Original Research Article

Received 20 May 2021
Accepted 25 July 2021
Published 19 August 2021

ABSTRACT

Aims: The main purpose of this study is to understand whether Logistic regression has certain benefits in the evaluation of American options. As far as the Monte Carlo method is concerned, the least square method is traditionally used to evaluate American options, but in fact, Logistic regression is generally quite good in classification performance. Therefore, this study wants to know if Logistic regression can improve the accuracy of evaluation in American options.
Study Design: The selection of options parameters required in the simulation process mainly considers the average level of actual market conditions in the past few years in terms of dividend yield and risk-free interest rate. The part of the stock price and the strike price mainly considers three different situations: in-the-money, out-of-the-money and at the money.
Methodology: This study applied the Logistic regression in Monte Carlo method for the pricing of American. Uses the ability of logistic regression to help determine whether the American option should be exercised early for each stock price path. The validity of the proposed method is supported by some vanilla put cases testing. The parameters used in all cases tested are considered the current state of the market.
Conclusion: This study demonstrates the effectiveness of the proposed approach using numerical examples, revealing significant improvements in numerical efficiency and accuracy. Several test cases showed that the relative error of all tests are below 1%.

**Corresponding author: E-mail: 138332@mail.tku.edu.tw;*

Keywords: American options; logistic regression; option pricing; Monte Carlo.

1. INTRODUCTION

An American option is a kind of an options contract that allows the holder to exercise the option at any time before and including the maturity day. Because of this feature of early exercise, it is necessary to compare the value of continuous holding and early exercise in the evaluation, and then decide whether to perform early. Therefore, so far, we have not found an analytical solution for its evaluation.

Due to the lack of analytic solutions to American options prices, researchers have developed a number of methods for these pricing problems. These methods include tree method, finite difference method and Monte Carlo simulation. In general, the advantages of the tree method and the finite difference method are that the calculation is fast and accurate, but the disadvantage is that they are not suitable for multi-assets and most of the exotic options. The advantage of Monte Carlo simulation is that it is more flexible and suitable for multi-asset and exotic options' evaluation. However, its disadvantage is that the computational complexity (complex in time and space) is relatively high.

Early studies using Monte Carlo simulation to evaluate American options include：Boyle (1977; for European option pricing) [1], Tilley (1993; American option pricing) [2], Barraquant and Martineau (1995; American option pricing) [3], Broadie and Glasser (1997; American option pricing) [4], Boyle, Broadie, and Glasserman (1997; American option pricing) [5] and Longstaff and Schwartz (2001; American option pricing) [6], Reesor and McLeish (2007; algorithom improvment) [7], Feng and Lin (2013; applied Levy Process Model in simulation) [8], Yari, G., M. Rahimi, P. Kumar (2017; multi-period multi-criteria) [9]. Among those researches, the Longstaff and Schwartz's method [LSM, Least Square Method] is perhaps the most popular and promising one of these methods. Many researchers have adopted, modified, and extended this method over the years LSM uses the regression method to predict the continuation value of each path, and compares it with the value of immediate exercise, and then decides whether to exercise it early.

1.1 Logistic Regression Application

Logistic regression is a technique that borrowed by machine learning from the field of statistics, it is used for the classification problems (such as whether to excercsie the American options or not is a classification problems), it is a predictive analysis algorithm and based on the theory of probability.

The key point for using Monte Carlo evaluate the price of American options is making a decision on whether to early exercise the option at each step. From this, using numerical pricing method can be seen as dealing with a classification problem. As long as you can correctly decide whether to excercise early, then you can accurately calculate the theoretical price of the American option.

Logistic regression is a discrete choice approach and suitable for useing as a tool for classification. So far, there are many finance or business related researches that are applied Logistic regression as following

1.2 Credit Risk

In the part of credit risk, it is mainly used to predict the possibility of a company's default in the future. In recent years, there have been many related studies as follows:

Deni (2015; assessing credit defaul) [10], Loredana and. Brédart. (2016; bankruptcy prediction) [11], Omar, Suwaidi. and Darshini Pun Thapa. (2012; credit risk management) [12], Ong, Yap and K.ong, (2011; corporate failure prediction) [13], Liou (2008; business failure prediction) [14], Spathis, C.T. (2002; detecting false financial statements) [15].

1.3 Other Business Issue

In addition to credit risk issues, logistic regression is also used in other management, including: Richard (2013; dividend policy and financial crisis) [16], Hiebl, Gärtner and Duller (2017; CFO's characteristics and ERP system adoption) [17], Kittilaksanawong (2014; issues on insider, and institutional shareholder) [18].

2. METHODOLOGY

2.1 Monte Carlo Algorithm

1. Generat N paths of stock prices, where the path *i=1,…..,M* evolves in discrete time

with index *j=1,…,N* (time interval $\Delta t = \frac{T}{N}$). The stocks process as following: $S_{i,j} = S_{i-1,j} e^{\left[\left(r-q-\frac{\sigma^2}{2}\right)\Delta t + \sigma\varepsilon\sqrt{\Delta t}\right]}$, r is the risk free interest rate, q is the dividend rate, σ is the volatility of the return of the stock and $\varepsilon \sim N(0,1)$.

2. In case of put option, at $t = N-1$, the present values of holding value are $e^{-rt}\max(K - S_{N,i}, 0)$ and the early exercise values are $\max(K - S_{N-1,i}, 0)$ (Since we cannot foresee the stock price of next day, so we don't have the exactly holding values at $N-1$, so we have to use logistic regression to predict which paths can be exercise earlier.)
3. According to the previous step, we took a peek at the stock price of the next period. Although we cannot directly use it to make decisions, we can set the depend variable to be 1 in the regression formula for the path that should be excercise early, and set the rest to be -1. The stock price is an independent variable to run logistic regression.
4. According to the result of step 3, the predicted value is used to decide whether to excercise early.
5. After completing step 4, we can get the holding value of $t = N-1$.
6. Next, go back to step 2 and calculate $t = N-2$ until $t = 1$ to get the price of the option.

2.2 Logistic Regression

The Logit model was developed by Berkson [19] and is a logistic regression model for solving problems of dichotomization. If we assume that the situation of early exercise or not at each stock price path is fi = (1: early exercise; -1: not early exercise), and let P be the probability of each path and S is the simulate stock price, we come up with the following equations:

$$P = E(f_i = -1 \mid S_i) \tag{1}$$

$$\text{Odds ratio} = \frac{P}{1-P} = \frac{\text{probability of early exercise}}{\text{probability of continuation}} \tag{2}$$

$$\begin{aligned} Ln\frac{P}{1-P} &= f(S_i\text{: simulated stock price}) \\ &= \beta_0 + \beta_1 S \cdots + \beta_k S^k + \varepsilon \\ P &= \frac{1}{1 + e^{\beta_0 + \beta_1 S \cdots + \beta_k S^k}} \end{aligned} \tag{3}$$

2.3 Application to the American Options

In this section, we are going to test the feasibility of logistic regression with some vallina cases. The cases we mainly discuss will be covered by some American Puts with several different parameter conditions.

First of all, we must decide all parameters of the examples. In order to improve the practicability of this method, this study tried to set the main parameters as close as possible to the real situation in the market. In the part of risk-free interest rate and diveden yield, this study is determined by reference to the actual situation of the current market. According to many reports and research institutions' information, the average dividend and net buyback yields was about 3.2% for those stocks are traded in S&P500 over the past few decades (Fig.1). With reference to this as a benchmark, the study set the ridk free interest rate q to be 3.5%.

Since the outbreak of credit crunch in 2008, the interest rate r (from LIBOR curve) has become much lower than earlier years and is quite close to zero in many countries. Fig. 2 shows how the interest rate r has changed over the period from 1985 to 2018.

According to the information, the average interest rate rate (one year LIBOR rate) was about 0.6% in the past ten years, and the interest rates rose slightly in the last 3 years, and it reached to an average about 1.7% in the past two years. Therefore, the study sets the risk-free interest rate levels to be 1% or 2% for all testing cases. Let it be closer to the real market situation.

In order to understand the validity of logistic regression, we will use Binomial tree as the benchmark, where the initial stock price S = 48(out of money), 50(at the money), and 52(out of money), time to expiration T = 1 year, volatility σ = 0.3, strike price K = 50. The other parameters are set as we mentioned in preview section.

In Tables 1 and 2 we present the results of the logistic regression, against the binomial method results obtained from the CRR tree with 10,000 time steps.

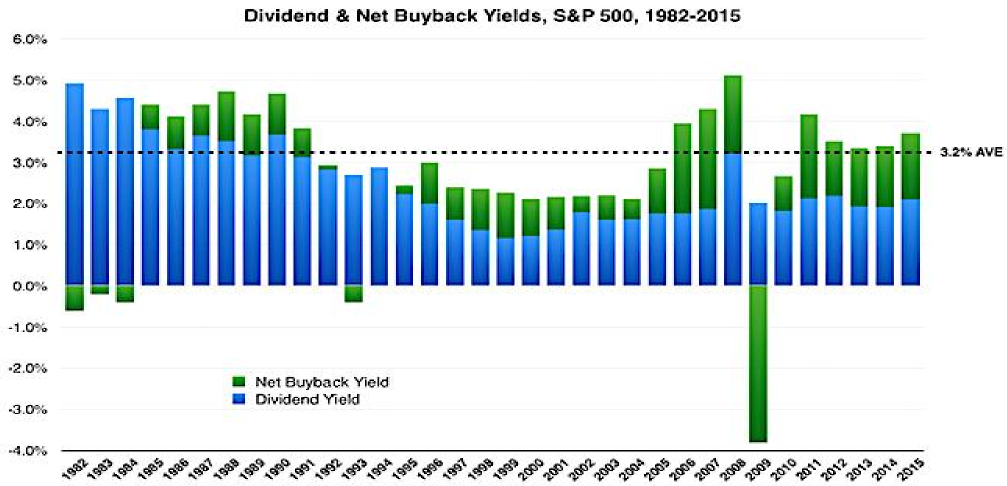

Fig. 1. Dividend & Buyback Yields, S&P 500, 1982-2015
Source: https://www.bogleheads.org/forum/viewtopic.php?t=193674

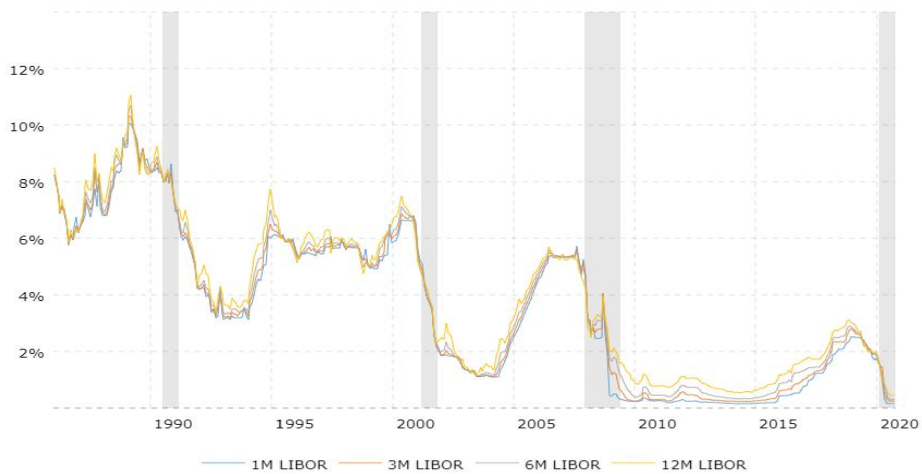

Fig. 2. The LIBOR rate in one year, 1985-2018
Source: https://www.macrotrends.net/1433/historical-libor-rates-chart

Table 1. American Vanilla Put Options

	CRR	Lgistic	RE
	r=0.01 (risk-free interest rate)		
S=48	7.4270	7.4168	-0.14%
S=50	6.4606	6.4564	-0.07%
S=52	5.5969	5.5911	-0.10%

Table 2. American Vanilla Put Options

	CRR	Lgistic	RE
	r=0.02 (risk-free interest rate)		
S=48	7.1143	7.1035	-0.15%
S=50	6.1731	6.1647	-0.14%
S=52	5.3345	5.3154	-0.36%

The accuracy is measured through relative error (RE).

$$RE_i = \frac{Logistic - CRR}{CRR} \tag{4}$$

where i means on different prices of underlying asset, Logistic is the price from our model and CRR is the price form benchmark.

3. CONCLUSION

The numerical results show that all of the relative error of the evaluation results using logistic regression are below 0.5% regardless of the stock prices are in or out of money, and it is obvious that the evaluation result has reliability. Longstaff and Schwartz [6] showed the RE in their research are between 0.89% and 0.00%. Compared with this research, the method of this research is obviously no less than the method of the above-mentioned literature.

Overall, we initially verified that logistic regression can be used for basic American options, and in the future we hope to further apply this method to other exotic options, including rain bow options and Asian options.

ACKNOWLEDGMENT

The author acknowledges the support from the Ministry of Science and Technology of Taiwan under the grant number 108-2410-H-432-001.

COMPETING INTERESTS

Author has declared that no competing interests exist.

REFERENCES

1. Boyle PP. “Options: A Monte Carlo approach. Journal of Financial Economics. 1977;4:328–338.
2. Tilley JA. Regression methods for pricing complex American-style options. Transactions of the Society of Actuaries. 1993;45:249–266.
3. Barraquant J, Martineau D. Numerical valuation of high dimensional multivariate American securities. Journal of Financial and Quantitative Analysis. 1995;30:383–405.
4. Broadie M, Glasserman P. Pricing American-style securities using simulation. Journal of Economic Dynamics and Control. 1997;21:1323–1352.
5. Boyle PP, Broadie M, Glasserman P. Monte Carlo methods for security pricing,” Journal of Economic Dynamics and Control. 1997;21:1267–1321.
6. Longstaff FA, Schwartz ES. Valuing American options by simulation: A simple least-squares approach. The Review of Financial Studies. 2001;14:113–147.
7. Reesor RM, McLeish D. A calibration algorithm for simulation-based pricing models. IMA Journal of Management Mathematics. 2007;18(4)371-393.
8. Lin LX, Feng. Pricing Bermudan Options in Levy Process Models,” SIAM Journal on Financial Mathematics. 2013;4 (1):474-493.
9. Gholamhossein Y, Mohamadtaghi R, Kumar Entropy P. Multi-period multi-criteria (mpmc) valuation of american options based on entropy optimization principles. Journal of Science and Technology. 2017;41(1):81-86.
10. Deni M. Assessing Credit Default Using Logistic Regression and Multiple Discriminant Analysis: Empirical Evidence from Bosnia and Herzegovina,” Interdisciplinary Description of Complex Systems. 2015;13(1):128-153.
11. Loredana C, Brédart X. Bankruptcy prediction: The case of Belgian SMEs. Review of Accounting and Finance. 2016;15(1):101-119.
12. Omar M, Al Suwaidi H, Darshini Pun Thapa P. Credit risk management: A case differentiating Islamic and non-Islamic banks in UAE", Qualitative Research in Financial Markets, 2012;4(2/3):197- 205.
13. Ong SW, Yap VC, K.ong RWL. Corporate failure prediction: a study of public listed companies in Malaysia. Managerial Finance. 2011;37 (6):553-564.
14. Liou FM. Fraudulent financial reporting detection and business failure prediction models: a comparison", Managerial Auditing Journal. 2008;3(7):650-66.
15. Spathis CT. Detecting false financial statements using published data: some evidence from Greece. Managerial Auditing Journal. 2002;17(4):179-191.
16. Richard H. Did dividend policy change during the financial crisis?. Managerial Finance, 2013;39(6):584-606.
17. Martin RW, Hiebl Gärtner, Duller B. Chief financial officer (CFO) characteristics and ERP system adoption: An upper-echelons

perspective. Journal of Accounting & Organizational Change. 2017;13(1):85-111.
18. Kittilaksanawong W. How do family, insider, and institutional shareholder perceive institutional risks in foreign market entry? evidence from newly industrialized economy firms. International Finance Review. 2014;15:279 – 306.
19. Berkson J. Application to the logistic function to Bio-Assay. Journal of the American Statistical Association. 1944;39:357-365.

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle4.com/review-history/70962