Journal of Business and Economics, ISSN 2155-7950, USA
November 2016, Volume 7, No. 11, pp. 1756-1763
DOI: 10.15341/jbe(2155-7950)/11.07.2016/004

http://www.academicstar.us

Pricing American Option via the Transform-Expand-Sample Forecasting Methods

AiChih Chang[1], Jim (Junmin) Shi[2]

(1. Rutgers University, USA; 2. New Jersey Institute of Technology, University Heights, USA)

Abstract: Pricing American options is notoriously intractable or computationally challenging due to its complexity feature of dynamic exercise strategy and the inherent uncertainty pertaining to its underlying asset price. Numerically, this study leverages the Least Squares method to price American options based on two simulation methods. In particular, to simulate the price process of the underlying asset, we propose the *Transform-Expand-Sample* (TES) approach, and compare its performance with the benchmark model of random walk. Random walk method is widely used if the volatility of the underlying asset price is the only factor affecting its behavior. In contrast, the TES approach is a versatile methodology for modeling stationary time series, whose principal merit is its ability to simultaneously capture first-order (marginal distribution) and second-order (autocorrelations) statistics of empirical time series. We experiment with several real-market American call options to illustrate the implementation of those two models. With an acceptable accuracy, the estimated option prices obtained by both approaches match the actual market price of the American option.

Key words: American option pricing; simulation; TES

JEL code: C630

1. Introduction

American options (*call* or *put*) are characterized by their dynamic exercising strategy and it is notoriously challenging to price their price even with a large-scale computational simulation. There are several algorithms for pricing American options leveraging Monte Carlo simulations (Wilmott P., 2006; Hull J. C., 1989). One of these most popular and practical approaches is so-called Least Squares approach, introduced by Longstaff and Schwartz (2011). In particular, the Least Squares method first simulates a number of the price paths of the underlying asset till its maturity time. Then for simulated each path, it calculates the discounted cash flows over time. The most critical step is that, at each time step, one compares the benefits of exercising decision with non-exercising the option in a simple regression manner.

To simulate the price paths of the underlying asset involves forecasting techniques for time series. In particular, if the risk-free rate and the volatility of the asset price are known, the Monte Carlo method could simulate future realizations by means of lognormal random walking. However, in most practical cases, the

AiChih Chang, CPIM, Rutgers Business School, Rutgers University; research areas/interests: supply chain management. Email: aichih@scarletmail.rutgers.edu.

Jim (Junmin) Shi, Assistant Professor, Tuchman School of Management, New Jersey Institute of Technology, University Heights; research areas/interests: supply chain management and finance. E-mail: Junmin.Shi@njit.edu.

volatility is not easy to be observed or is significantly time varying. Based on empirical data, to make accurate forecasting becomes overwhelmingly critical. It is well known that the observed empirical data contains a great deal of meaningful information. Statistically, the marginal distribution and autocorrelations are the main statistics which can be obtained directly from the empirical data and thus play an important role in forecasting models. From the simulation perspective, the histogram inversion can be utilized to generate random variables which have the same density as the one pertaining to the empirical data. Furthermore, since most of the financial time series are highly auto-correlated, the simulated data should have the same autocorrelation statistics as that of empirical data. To this end, the *Transform-Expand-Sample* (TES) methodology was introduced in the early 1990's (Jagerman D. L. & Melamed B., 1992). The principal merit of TES approach is its ability to simultaneously capture first-order (marginal distribution) and second-order (autocorrelations) statistics of empirical time series. As the original purpose, it was designed to model telecommunication traffic, especially in emerging high-speed communications networks. In this study, we shall apply the TES methodology to forecast the underlying asset price and then leverage the Least Squares approach to price American options.

2. Pricing American Option

2.1 Data Collection

We collected the weekly index of S&P 100 from May 5, 2000 to May 4, 2007, and the corresponding S&P index call option prices with various maturity times: May 18 (2 weeks), Jun. 15 (6 weeks), Jul. 20 (11 weeks), Aug. 17 (15 weeks), Sep. 21 (20 weeks), and Dec. 21 (33 weeks). We use the Treasure Bill (T-bill in short) as the risk free asset. According to the public data of T-bill, its return rates of 4 weeks, 13 weeks and 26 weeks are 4.61%, 4.79%, and 4.82%, respectively. The T-bill return rates of other length of weeks are obtained by a linear-piecewise function of the aforementioned 4 weeks, 13 weeks, and 26 weeks. For example, the T-bill return rate of 6 weeks is approximated as:

$$r = \frac{6-4}{13-4} \times 4.61\% + \frac{13-6}{13-4} \times 4.79\% = 4.75\% \cdot$$

2.2 Simulating the Asset Price Path

2.2.1 Monte Carlo Method of Random Walk

From the S&P 100 weekly index, we estimate the volatility of the weekly return as σ_w^2 from which we estimate the annual volatility as $\sigma^2 = 52\sigma_w^2$. Then the future's weekly price of the underlying asset is obtained iteratively by

$$S_{j+1} = S_j \exp\left((r - \frac{\sigma^2}{2}) \delta t + \sigma \sqrt{\delta t} Z \right) \tag{1}$$

Where δt = 1/52 and the noise term Z follows a standard normal distribution, i.e., $Z \sim N(0,1)$.

2.2.2 TES (Transform-Expand-Sample)

The empirical marginal density of an empirical sample is estimated in practice by an empirical histogram statistic in the form of $\hat{H} = \{(l_j, r_j, \hat{p}_j) : 1 \le j \le J\}$, where J denotes the number of histogram cells, $[l_j, r_j)$ is the j_{th} cell with width $w_j = r_j - l_j$, and $\hat{p}_j$ is the probability estimator for the j_{th} cell. The empirical *probability density function* (pdf) is estimated as

$$\hat{h}(y) = \sum_{j=1}^{J} 1_{[l_j, r_j)}(y) \frac{\hat{p}_j}{w_j}, \quad -\infty < y < \infty \tag{2}$$

Where indicator function $1_A(x)$ is 1 if $x \in A$; 0 otherwise. The corresponding *cumulative distribution function* (cdf) is a piecewise linear function given as

$$\hat{F}(y) = \sum_{j=1}^{J} 1_{[l_j, r_j)}(y) \left[\hat{C}_{j-1} + (y - l_j) \frac{\hat{p}_j}{w_j} \right], \quad -\infty < y < \infty \tag{3}$$

Where $\hat{C}_j = \sum_{i=1}^{j} \hat{p}_i$ for $1 \le j \le J$, $\hat{C}_0 = 0$ and $\hat{C}_J = 1$.

The *histogram inversion* is a piecewise linear function as follows

$$D_H(x) = \hat{F}^{-1}(x) = \sum_{j=1}^{J} 1_{[\hat{C}_{j-1},\, \hat{C}_j)}(x) \left[l_j + (x - \hat{C}_{j-1}) \frac{w_j}{\hat{p}_j} \right], \quad 0 \le x \le 1 \tag{4}$$

For a discrete-time stationary stochastic process, $\{X_n\}_{n=0}^{\infty}$, its autocorrelation function $\rho_X(\tau)$ consists of the lagged correlation coefficients,

$$\rho_X(\tau) = \frac{E[X_n \ X_{n+\tau}] - \mu_X^2}{\sigma_X^2}, \tag{5}$$

Where $\mu_X < \infty$ and $\sigma_X^2 < \infty$ are the common mean and variance, respectively.

The TES model provides a generic scheme for generating stationary sequences.

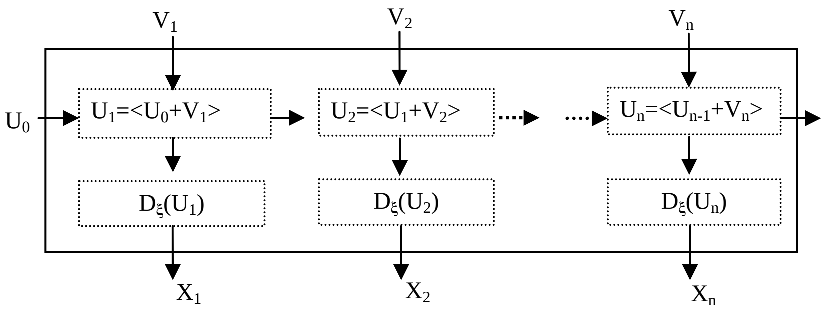

Figure 1 Foreground/Background Schemes

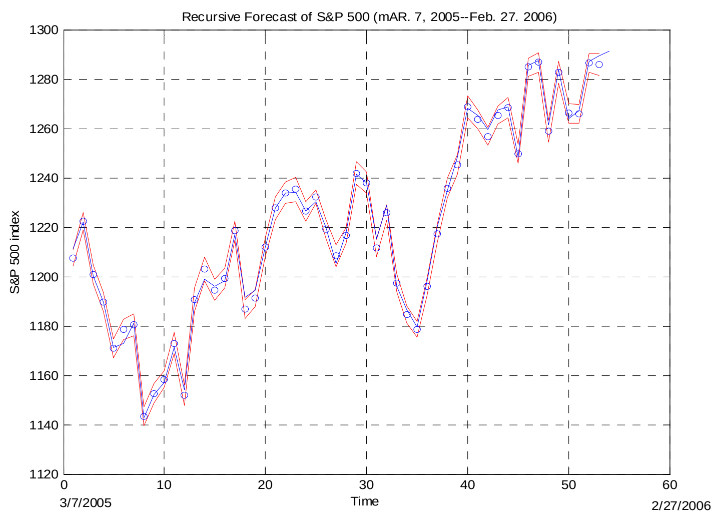

Figure 2 S&P 500. (3/7/2005-2/27/2006) and TES Recursive Weekly-Forecast

Shown in Figure 1 is the generic sketch of TES process, where $\boldsymbol{D}$ (referred to as the *distortion*) transforms the background process $\{U_n\}$ to a corresponding foreground process $\{\boldsymbol{X}_n\}$. In particular,

(1) $\{\boldsymbol{V}_n\}$ is a sequence of innovations (i.i.d random variables) with common density $\boldsymbol{f}_V$, independent of $\boldsymbol{U}_0$. In particular, $\boldsymbol{f}_V$ is approached by *step-function*

$$f_V(x) = \sum_{k=1}^{K} 1_{[L_k,R_k)}(x)\frac{P_k}{R_k - L_k}, \qquad (-0.5 \le x \le 0.5) \tag{6}$$

Where $\boldsymbol{K}$ is the number of steps, $\boldsymbol{P}_k$ is the mixing probability of step $\boldsymbol{k}$.

(2) $U_0 \sim Uniform(0,1)$ is the initial variant in a background TES sequence.

(3) The *background sequence:* $U_n = \begin{cases} U_0 & n = 0 \\ < U_{n-1} + V_n > & n > 0 \end{cases}$.

(4) $D = F_X^{-1}(S_\xi(x))$ is a distortion function, in which $\boldsymbol{S}_\xi$ is defined by a *stitching parameter* ξ

$$S_\xi(y) = \begin{cases} y / \xi & 0 \le y \le \xi \\ (1-y) / (1-\xi) & \xi \le y \le 1 \end{cases}. \tag{7}$$

(5) The foreground sequence $\{\boldsymbol{X}_n\}$ is obtained based on corresponding background sequence $\{U_n\}$,

$$X_n = D(U_n) \tag{8}$$

Mathematically, for a given lag τ, the corresponding autocorrelation is

$$\rho_X(\tau) = \frac{2}{\sigma_X^2}\sum_{v=1}^{\infty} \Re[\hat{f}_V^\tau(i2\pi v)]\left|\hat{D}(i2\pi v)\right|^2 \tag{9}$$

Where $\hat{f}_V(s) = \int_{-\infty}^{\infty} e^{-sx} f_V(x)dx$ is the Laplace transform of the pdf of innovation $\{V_n\}$.

Typically, the following outlines the algorithm for the TES method.

(a) Setting up the empirical histogram inversion and autocorrelation function: The histogram inversion and autocorrelation function are computed from empirical data,

(b) Selecting the stitching parameter and innovation density: The core activity is to find the optimal innovation density f_V and stitching parameter ξ to minimize the weighted difference between empirical autocorrelation in Equation (5) and the TES autocorrelation in Equation (9).

(c) Generating TES sequences: With the optimal parameters f_V and ξ obtained in (b), we generate the TES sequences according to the scheme depicted in Figure 1.

Figure 2 depicts the results of weekly forecasting by TES. Graphically, TES performs relatively well on one-week forecasting.

2.3 Pricing the American Option

2.3.1 Example for the Least Squares method

To illustrate the algorithm in detail, we focus on a simple example for pricing the America Call option. First, we generate 5 realizations of the asset path from now on to its maturity date. The initial price of the underlying stock is $\boldsymbol{s}_0 = 689.4$, the strike price is $k = 670$, and the risk free rate be $r = 0.0457$. Note that the discount coefficient can be computed by $\beta = e^{-r/52}$.

Table 1 Eight Paths Simulated by Random Walk Method

	S_0	S_1	S_2	S_3	Payoff
Path 1:	686.4	709.93	717.1	714.3	44.2951
Path 2:	686.4	695.03	689.49	702.59	32.5902
Path 3:	686.4	708.97	715.25	709.23	39.2334
Path 4:	686.4	686.52	676.11	654.05	0
Path 5:	686.4	678.55	679.93	668.33	0
Path 6:	686.4	687.81	668.49	679.4	9.3975
Path 7:	686.4	702.02	697.53	696.91	26.9139
Path 8:	686.4	678.7	683.9	688.23	18.2317

Table 2 First Step Regression

	Y	X
Path 1:	ß *44.2951	717.1
Path 2:	ß *32.5902	689.49
Path 3:	ß *39.2334	715.25
Path 4:	ß *0	676.11
Path 5:	ß *0	679.93
Path 6:		
Path 7:	ß *26.9139	697.53
Path 8:	ß *18.2317	683.9

The computational mechanics is described as follows:

(1) Calculating the payoffs at the expiration for each path (please refer to Table 1). Table 1 depicts the expected value what we would receive if we exercised on the maturity.

(2) Determining whether to exercise earlier. Going backward to time 2, we calculate two column values: X (the stock price at time 2, for those paths which are in the money at that time) and Y (the payoffs for these paths, discounted back from maturity to time 2); cf. Table 2.

(3) Calculating the cash flow from holding the option, conditional on the stock price at time 2. Use a regression of Y as a quadratic function of X by least squares methods: $Y = 0X^2 + 44X - 15562$.

(4) Comparing the two values, “Exercise now” and the value computed by the regression function, and then selecting the one with higher value; cf. Tables 3 & 4.

Table 3 Regression

	Exercise now	by the reg.
Path 1	47.1	41.2256
Path 2	19.49	23.9502
Path 3	45.25	41.5287
Path 4	6.11	-1.2311
Path 5	9.93	7.0774

Table 4 Cash Flows

S_2	S_3
47.0957	0
0	32.5902
45.2461	0
6.1080	0
9.9267	0

Table 5 Final Cash Flows

S_1	S_2	S_3
0	47.0957	0
25.0324	0	0
0	45.2461	0
16.5191	0	0
0	9.9267	0

(5) By repeating the same calculation backward to time zero, we obtain the final cash flow matrix; cf. Table 5.

(6) The last step is to present the values for each of these cash flows backward to time zero and average them as the option price. In this example, the option price is computed to be $26.448.

2.3.2 Implementations

For a given American call option, we simulate 100 paths by random walk method and TES. With the simulated 100 paths, we compute the option price via least squares method. By repeating the same experiments for 100 times, we take the mean value as the option price and obtain the 95% confidence intervals via computing z-test.

3. Simulation Performance

3.1 Option Pricing Based on Random Walk Simulation

The first three sub-pictures in Figure 3 depict the computed option price compared with its market prices. Each of them has five curves: red curve (bid price), green curve (asked price), black curve (estimated price), and two yellow dash-dot curves (the upper and lower boundaries for the 95% confidence interval). It is shown that the computed option prices via the random walk simulation are relatively approximating the market prices. The forth sub-picture depicts the estimated prices for American calls with different maturities. For the same strike price, it is shown that the option price increases as the option lifetime gets longer.

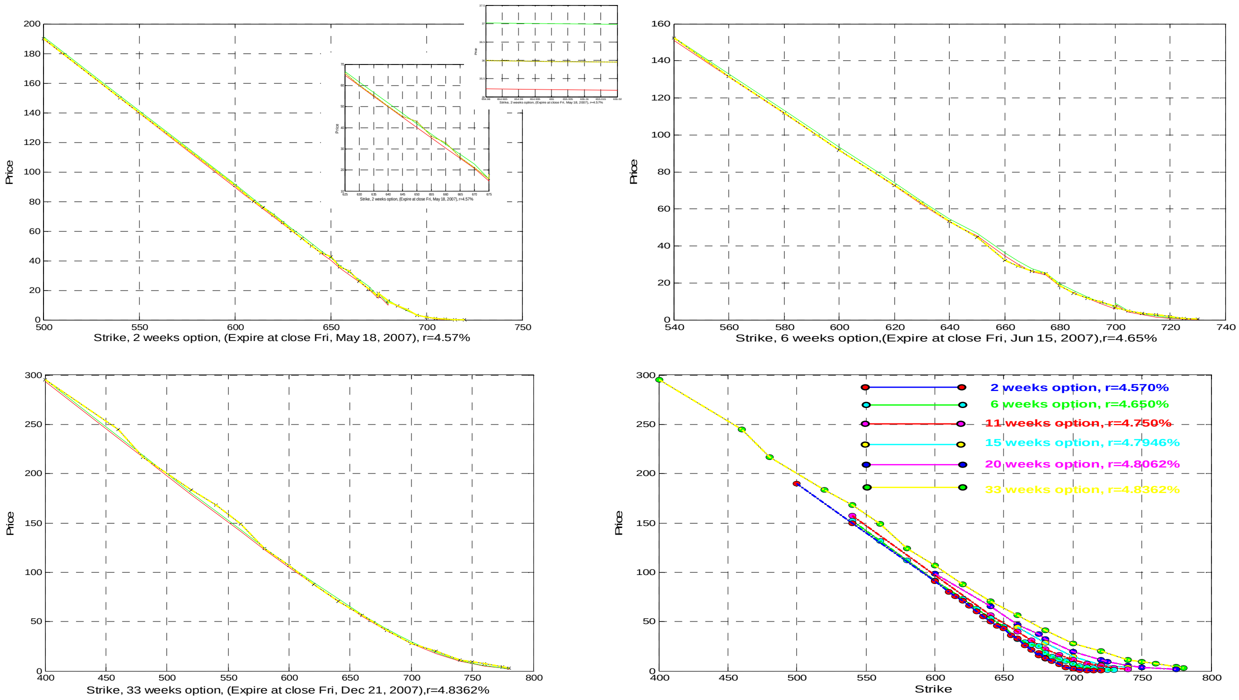

Figure 3　Computed Option Prices (Random Walk) v.s. Market Prices: Selected Maturity for Varying Strike Prices

3.2 Performance of the TES Method

Leveraging the TES approach to simulate the asset price paths instead of random walk approach, we obtained the call price with the Least Squares method in option pricing. The following figures depict the estimated option prices via the TES method for different maturities.

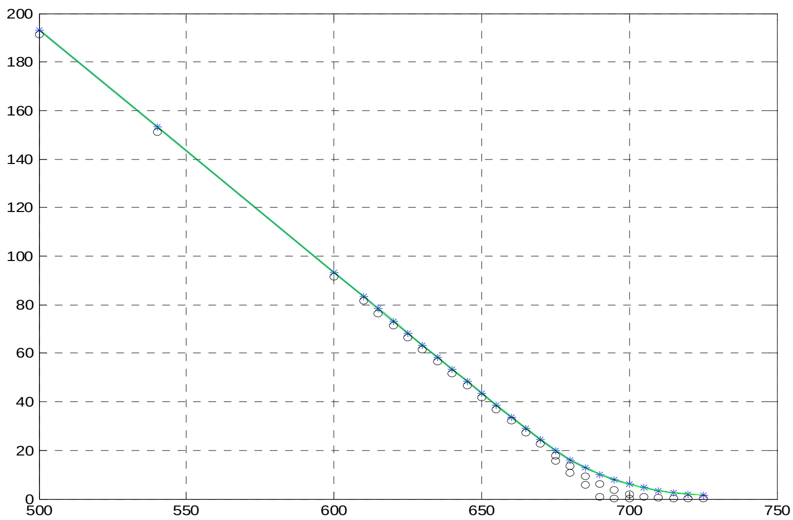

Figure 4(a)　2 Week Call of S&P 100

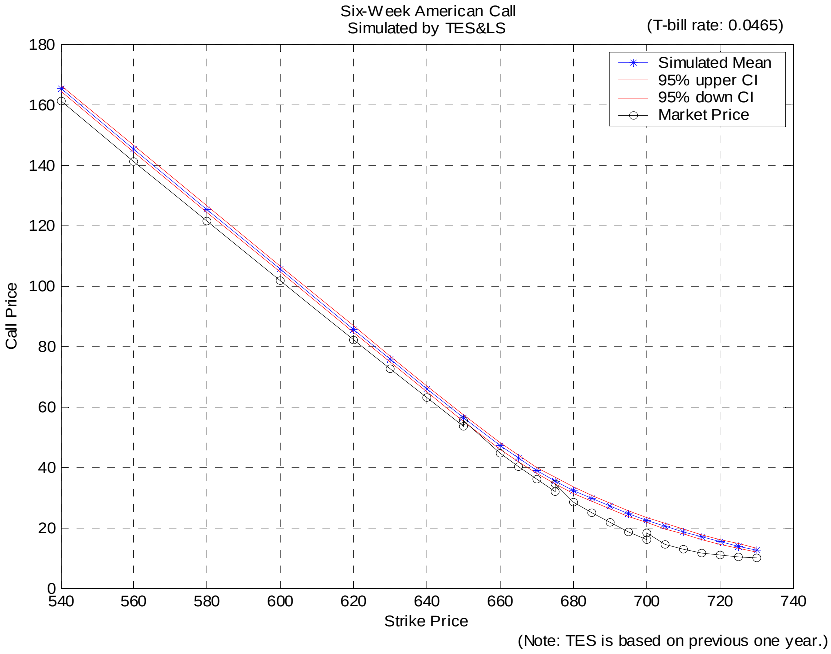

Figure 4(b)   6 Week Call of S&P 100

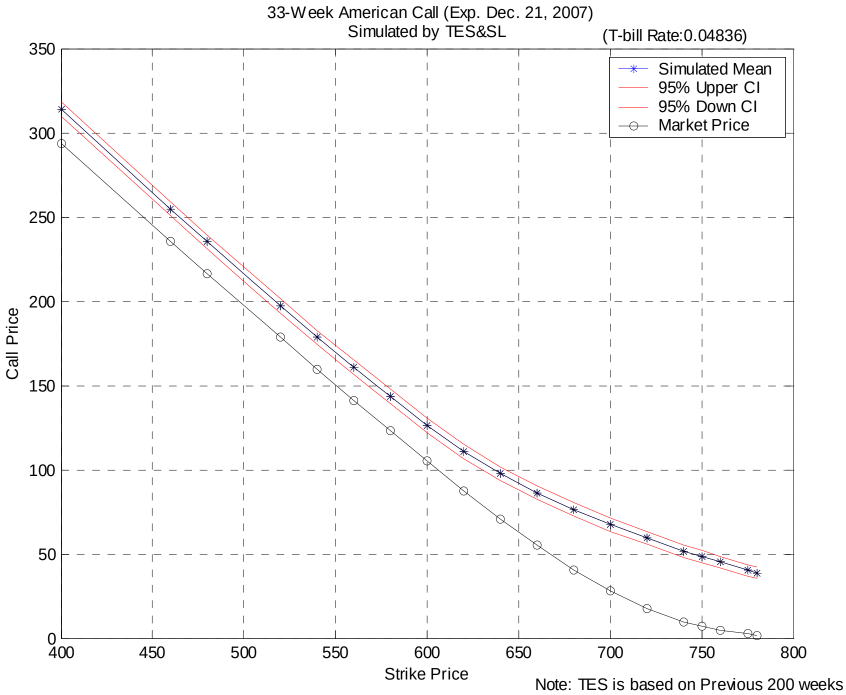

Figure 4(c)   33 Week Call of S&P 100

Figure 4(a) depicts the computed price for a call with two week maturity of varying strike prices. In a similar vein, Figures 4(b) and 4(c) display the corresponding results for options of 6 weeks and 33 weeks, respectively. In each of them, the blue curve indicated with (*) is the mean price computed with 100 iterations, the area between the red curves is the 95% confidence intervals, and the black curve indicated with (o) is the market price. Roughly speaking, the gap between estimated price and market price grows larger as the maturity of the option gets longer.

Furthermore, for a given option the confidence intervals get wider when the option maturity gets longer. Another observation is that the TES approach performs well in a short term forecasting.

4. Conclusion

We leveraged the Least Squares method to price American options. To simulate the asset price, we compared random walk and TES approaches. In general, the Least Squares random walk simulation could fit a good model to the option market. In contrast, the TES performs well in short term forecasting but gets poor for long ones. For each of the two approaches, the confidence interval gets wider as the option maturity gets longer. One of the intuitive reasons is that the option market accumulates uncertainty over time. One possible explanation for the error of both methods is the granularity of time intervals: the granularities in the models are weeks instead of days.

References

Hull J. C. (1989). *Options, Futures, and Other Derivative Securities* (5th ed.), Prentice-Hall, pp. 474-477.

Jagerman D. L. and Melamed B. (1992). "The transition and autocorrelation structure of TES processes, Part I: General theory", *Stochastic Models*, Vol. 8, No. 2, pp. 193-219.

Jagerman D. L. and Melamed B. (1992). "The transition and autocorrelation structure of TES processes, Part II: Special cases", *Stochastic Models*, Vol. 8, No. 3, pp. 499-527.

Longstaff F. A. and Schwartz E. S. (2011). "Valuing American options by simulation: A simple least-squares approach", *Review of Finance Studies*, Vol. 14, No. 1, pp. 113-147.

Wilmott P. (2006). *Paul Wilmott on Quantitative Finance* (2nd ed.), Vol. 3, pp. 1279-1283.