# AN IMPROVED APPROACH FOR VALUING AMERICAN OPTIONS AND THEIR GREEKS BY LEAST-SQUARES MONTE CARLO SIMULATION

YOUNGSOO CHOI
DEPARTMENT OF MATHEMATICS
HANKUK UNIVERSITY OF FOREIGN STUDIES
YONGIN-SHI, KYONGKI-DO
KOREA, 449-791

JOONHYUK SONG
KOREA DEVELOPMENT INSTITUTE
DONGDAEMUN-GU, SEOUL
KOREA, 130-012

# AN IMPROVED APPROACH FOR VALUING AMERICAN OPTIONS AND THEIR GREEKS BY LEAST-SQUARES MONTE CARLO SIMULATION

YOUNGSOO CHOI

DEPARTMENT OF MATHEMATICS

HANKUK UNIVERSITY OF FOREIGN STUDIES

YONGIN-SHI, KYONGKI-DO

KOREA, 449-791

JOONHYUK SONG

KOREA DEVELOPMENT INSTITUTE

DONGDAEMUN-GU, SEOUL

KOREA, 130-012

Abstract. This paper presents a new methodology to approximate the value of American options by least-squares Monte Carlo simulation. Whereas Longstaff and Schwartz's approach do not utilize the underlying asset price movement, we develop several methods that incorporates it into option pricing. One category improves the R-squares from the regressions by using, [1] the weighted regression with the same regressors and, [2] new regressors which are related to the discount factor from the current decision to exercise time. The other category improves the computational speed without sacrificing the convergence level by, [1] terminating early during the backwardation procedure and, [2] decreasing the number of observations for the regressors. Finally, combining both methods, we can get improved R-squares and computational speed in comparison to Longstaff and Schwartz's approach.

*JEL classification*: G10; G13

*Keywords*: least-squares Monte Carlo simulation; American put options; Hedge Parameters

---

*Date*: May 15, 2006 for this draft.

We are responsible for possible errors and omissions. The corresponding author's e-mail address is `choiys@hufs.ac.kr`.

## 1. Introduction

The valuation and properties of American options has drawn a lot of attention from both academics and industry because of its far-ranging applicability in financial markets, where a variety of financial instruments are traded including stocks, foreign currencies, interest rates, commodities, and future contracts. However, American options are typically very difficult to analyze due to the intrinsic uncertainty associated with the exercise rule which offers the holder the right to exercise the option early to liquidate the payoff before the expiration date. Therefore, one of many challenges in analyzing American options lies in finding the optimal exercise rule from the underlying asset's stochastic properties. Hence, quite often, the properties of American options are deduced from their European equivalent.

For the last decade or so, numerical procedures for valuing American options have been widely utilized to compute option prices using Monte Carlo simulation. In their seminal work, Cox, Ross and Rubinstein (1979) apply a binomial pricing model to solve the optimal exercise rule and the option price. The binomial model is still one of the most popular pricing formulae in practice. Despite its convenience in implementation and conceptual ease, the computational burden of the binomial model grows exponentially as the number of state variables increases. Many researchers have developed more advanced and complicated numerical algorithms than can be employed to price options since the pioneering work of Cox et. al. (1979), but a major drawback of those algorithms is their inaccuracy and slowness of computation.

A number of papers using simulation follows. Bossaerts (1989) solves the exercise rules which maximizes the simulated value of the option. Averbukh (1997), Broadie and Glasserman (1997), Carr (1998), and Garcia (1999) use various parameterization techniques to approximate exercise boundaries or transitional densities, to name a few.

Recently, Longstaff and Schwartz (2001) proposed a simple but powerful numerical procedure to compute option prices using a regression framework, which they call LSM (Least Squares Monte Carlo). They estimated the expected cash flows conditional on in-the-money stock prices and compare the expected payoff with the immediate exercise value of the option and obtained the optimal exercise rule only when the option value of the immediate exercise is higher than the expected payoff. This process continues backward until the starting date for each simulation and evaluates the option price by taking the average of the option prices of all simulation runs. One of the merits of Longstaff and Schwartz's approach is that it does not require a parametric exercise boundary. Their simple but powerful approach has ignited a renewed interest in regression-based American option pricing. Pizzi and Pellizzari (2002) use a non-parametric technique to tackle the problem, which gives the advantages of not needing to choose transformed regressors unlike the LSM. However, the non-parametric pricing method can be applied only when the number of underlying assets are few. The typical curse of dimensionality problem of non-parametric estimation directly constrains its applicability in a more general setting.

This paper aims to provide a few extensions to approximate the value of American options by least-squares Monte Carlo simulation. While Longstaff and Schwartz's method do not utilize the underlying asset price movement, we develop several methods that incorporates it into option pricing. One category improves the R-squares by using the weighted regression with the same regressors suggested by them and the other is related to the discount factor from the current decision to exercise time. The second method improves the computational speed without sacrificing the convergence level by terminating early during the backwardation process and decreasing the number of observations for the regressors. We combine both methods and achieve a more progress in R-squares and computational speed compare to Longstaff and Schwartz's approach.

The rest of the paper is organized as follows. Section 2 briefly reviews the working mechanism of LSM. Section 3 introduces two extensions to solve option pricing problems. Also, we present two methods that improve the computational speeds in pricing simulation. Section 4 discusses the sensitivity of the option to the change in underlying parameters, known as the *Greeks* and shows how the calculation of the *Greeks* can be achieved within our simulation schemes. A variety of simulation results will be presented and discussed in Section 5. Concluding remarks will follow.

## 2. Longstaff and Schwartz Model for Option Pricing

We consider the American put option on a no-dividend stock with strike price, $K$, and time-to-maturity, T. At the exercise time, $t_e(< T)$, the immediate exercise value, $V_{t_e}^e = \max(K - S_{t_e}, 0)$, is well-fitted through regression as follows:

$$V_{t_e}^e = \alpha + \beta_1 S_{t_e} + \beta_2 S_{t_e}^2 + \varepsilon. \tag{1}$$

According to the LSM approach based on Longstaff and Schwartz (2001), at the decision time, $t(< t_e < T)$, the optimal strategy is to compare the immediate exercise value, $V_t^e$, with the discounted cash flows from continuing, and then exercise if immediate exercise is more valuable. As a benchmark, we will recap the LSM approach and show the estimation methods applied in their approach.

The LSM approach provides the stopping rule that maximizes the value of the option at each decision time along each path and uses only in-the-money paths in the estimation, since the exercise decision is only relevant when the option is in the money.

For the computation of the discounted expected value of continuation at the decision time $t$ the LSM selects in-the-money paths, then moves forward along the selected path until the first stopping time $t_e$ occurs, and finally use the discounted cash flow from the exercise back to time $t$, for example, in case of the typical put option, $Y = e^{-r(t_e - t)} \max(K - S_{t_e}, 0)$, as the dependent variable. Also the stock price at time $t$, $S_t$, is used as the independent variable, $X$, for the following regression model[1]:

$$Y = \alpha + \beta_1 X + \beta_2 X^2 + \varepsilon. \tag{2}$$

[1]Longstaff and Schwartz argue that simple power functions of asset prices yield similar results as polynomial transformations of asset prices are used as basis functions.

Since the LSM does not take into account the relationship between the behavior of the stock price process and the selected regressor in the regression, the $R^2$s from the regressions are often somewhat low and thus makes the measurement error in the estimated continuation value a little too large. As alternatives, regressors taking into account the property of the stock price process will be introduced.

## 3. Improved Algorithms for Option Pricing

### 3.1. Improvement of $R^2$s from regressions.

*A. Weighted regression model*

For a simulated path $\omega_i$, $i = 1, \cdots, N$, from the decision time, $t$, until the first stopping time or exercise time, $t_e^{(i)}$, the stock price under the risk-neutral measure can be generated from the formula:

$$S_{t_e}^{(i)} = S_t^{(i)} \exp\big((r - \sigma^2/2)(t_e^{(i)} - t) + \sigma\sqrt{t_e^{(i)} - t}\, \phi_e^{(i)}\big), \tag{3}$$

where $S_t^{(i)}$ is the stock price at time $t$, $r$ is the riskfree interest rate, $\sigma$ is the stock's volatility, and $\phi_e^{(i)}$ is an independent sample from the standard normal distribution. Thus, the log value of the ratio of realized stock price at $t_e^{(i)}$ to stock price at $t$ follows a normal distribution:

$$\log\big(S_{t_e}^{(i)}/S_t^{(i)}\big) = (r - \sigma^2/2)(t_e^{(i)} - t) + \sigma\sqrt{t_e^{(i)} - t}\, \phi_e^{(i)}.$$

The realized normal shock follows a standard normal distribution:

$$\phi_e^{(i)} = \tfrac{1}{\sigma\sqrt{t_e^{(i)} - t}} \Big[ \log\big(S_{t_e}^{(i)}/S_t^{(i)}\big) - (r - \sigma^2/2)(t_e^{(i)} - t) \Big] \;\sim\; N(0,1).$$

Hence, the probability that the stock price $S_t$ at time $t$ moves to $S_{t_e}^{(i)}$ at time ${t_e}^{(i)}$ can be obtained from the probability density function as follows:

$$w_e^{(i)} = \tfrac{1}{\sqrt{2\pi}} e^{-{\phi_e^{(i)}}^2/2}.$$

The ordinary least squares estimates of parameters $\{\alpha, \beta_1, \beta_2\}$ in regression (2) is obtained by minimizing the object function:

$$\sum_{\{i\,|\,Y_i>0\}} w_e^{(i)}(Y_i - \alpha - \beta_1 X_i - \beta_2 X_i^2)^2, \tag{4}$$

where $Y_i = e^{-r(t_e^{(i)}-t)} \max(K - S_{t_e}^{(i)}, 0)$ and $X_i = S_t^{(i)}$.

*B. Regression with new regressors*

It is well known that the $R^2$s for regression (1) with regressors $1, S_{t_e}$, and $S_{t_e}^2$ are very high and the discounted expected cash flow of continuation should be regressed by the stock price at the decision time, $S_t$. Thus, substituting the regressor related to $S_{t^e}$ in equation (1) with the current time stock price $S_t$ through equation (3), we have a new regression model for the dependent variable $Y_i = e^{-r(t_e^{(i)}-t)} \max(K - S_{t_e}^{(i)}, 0)$ as follows:

$$Y_i = \alpha e^{-r(t_e^{(i)}-t)} + \beta_1 S_t^{(i)} e^{-\sigma^2(t_e^{(i)}-t)/2} + \beta_2 (S_t^{(i)})^2 e^{(r-\sigma^2)(t_e^{(i)}-t)} + \varepsilon_i, \tag{5}$$

where the error term $\varepsilon_i$ can be written:

$$\varepsilon_i = \varepsilon + \phi_e^{(i)}\, \sigma \sqrt{t_e^{(i)} - t}\, e^{-\sigma^2(t_e^{(i)}-t)/2}\, S_t^{(i)} \big(\beta_1 + 2\beta_2 S_t^{(i)} e^{(r-\sigma^2/2)(t_e^{(i)}-t)}\big).$$

That is, the new independent variables, $e^{-r(t_e^{(i)}-t)}$, $S_t^{(i)} e^{-\sigma^2(t_e^{(i)}-t)/2}$, and $(S_t^{(i)})^2 e^{(r-\sigma^2)(t_e^{(i)}-t)}$, related to the discount factor from the decision to first stopping time, are introduced to improve the $R^2$ values from the regressions.

Although the unspecific risk of equation (5) is increased in comparison to regression (1), the $R^2$s are substantially improved in comparison to the Longstaff and Schwartz model which will be shown in our simulation section.

3.2. **Improvement of computational speed.**

Due to the backward induction nature of pricing American options, most numerical algorithms require a large storage space. Thus, the computational speed heavily relies on the particular algorithm that is chosen.

In this subsection we introduce new methodologies to improve the computational speed without sacrificing the convergence level by, [1] terminating early during the backwardation procedure and [2] decreasing the number of observations for the regressors.

*C. Early termination Model*

Recall that the optimal exercise boundary at the decision time $t$ for the American put option with maturity $T$ decreases as the time-to-maturity $T-t$ increases. This phenomena is shown in Figure 1, where the benchmark optimal exercise boundary is obtained from the implicit finite difference method (IFDM). Also, it can be observable that the minimum value of simulated paths at each decision time, $t$, is almost increasing as the time-to-maturity increases, since the volatility of stock price at time $t$, $\sigma\sqrt{t}$, is decreasing as the time-to-maturity $T-t$ increases.

Combining these two observations, we suggest an *early termination algorithm* for three models described in the former subsection; For simplicity, we focus the discussion on the case where the American option can only be exercised at the $n$ discrete times $0 < t_1 < t_2 < \cdots < t_n = T$, and consider the optimal termination policy at each exercise time. During the backwardation procedure starting the final maturity time, $T$,

Step 1. Compute the optimal exercise boundary at the decision time, $t_k$, for the LSM approach by taking the maximum stock price among the pathes at which immediate exercise, $V^e_{t_k}(S^{(i)}_{t_k})$, is more valuable than the discounted expectation of continuation, $V^c_{t_k}(S^{(i)}_{t_k})$, as follows:

$$S^{(b)}_{t_k} = \max_{\{S^{(i)}_{t_k} \,|\, V^e_{t_k}(S^{(i)}_{t_k}) > V^c_{t_k}(S^{(i)}_{t_k})\}} S^{(i)}_{t_k}.$$

Step 2. If the minimum value of simulated paths at the decision time $t_k$ defined by

$$S_{t_k}^{(m)} = \min_{i=1,\cdots,N} S_{t_k}^{(i)},$$

equals to $S_{t_k}^{(b)}$, stop backwardation procedure and identify the cash flows at each path based on the information set at time $t_{k+1}$ instead of $t_k$. That is, it is more reasonable to use the information about exercise time $t_e^{(i)}$ for each path at time $t_{k+1}$, since the optimal exercise boundary at time $t_k$ hits the minimum value of simulated paths.

*D. Flexible upper bound for regressor's observations*

Note that the conditional expectation of stock price at maturity based on the information $\mathcal{F}_t$ at time $t$ is given by

$$\mathrm{E}^*[\, S_T \mid \mathcal{F}_t \,] = S_t e^{r(T-t)}.$$

where the asterisk indicates the expectation is taken in the risk-neutral measure. Since the optimal exercise boundary at the decision time, $t$, for the American put option with strike price, $K$, and maturity, $T$, is decreasing as the time-to-maturity, $T-t$, increases and its optimal exercise boundary at maturity is K, it is not necessary to include the full in-the-money paths as the independent variables in the linear regression model. Thus, we propose a new rule to improve the computational speed and get more accurate exercise boundary for the American put option by changing the selection methodology for the independent regressor as follows:

Step 1. Compute the cumulative probability that simulated paths at each decision time, $t$, are less than the time-adjusted strike price, $Ke^{-r(T-t)}$:

$$p_{t,T} = \mathbb{P}^*[\, S_t^{(i)} \mid S_t^{(i)} < Ke^{-r(T-t)} \,].$$

Step 2. Determine the upper bound for the independent regressor's observations in the linear regression model at each decision time $t$, which is $U_r(t) = K = 40$ for the fixed case like Longstaff and Schwartz's approach and

$$U_r(t) = S_0 e^{(r-\sigma^2/2)t+\sigma\sqrt{t}q_{t,T}}$$

for the flexible case, where $q_{t,T}$ satisfies $\mathbb{P}^*[\,\phi \le q_{t,T}\,] = \xi {p_{t,T}}^2$ with $\phi \sim N(0,1)$.

Step 3. For the parameter estimation of the regression model (2), (4), and (5) at the decision time, $t$, we use the simulated paths for the independent variable satisfying the inequality $S_t^{(i)} < U_r(t)$.

## 4. Hedge Parameters

In many cases, it is important to know the sensitivity of the option to the changes in the underlying assets of option parameters. These parameters are collectively referred to as *the Greeks*. Mathmatically speaking, the Greeks indicate the derivatives of a derivative security's price with respect to various model parameters. For each simulated path $\omega_i,\ i = 1, \cdots, N$, the LSM algorithm produces an exercise time $t_e^{(i)}$ and its discounted payoff function $\hat{P}_i(S_0) = e^{-rt_e^{(i)}} \max(K - S_{t_e^{(i)}}, 0)$ so that the American put option price, $P(S_0; K, T)$, is approximated by the sample average:

$$\hat{I}_N = \frac{1}{N}\sum_{i=1}^{N} \hat{P}_i(S_0).$$

The convergence result from Longstaff and Schwartz addresses the bias of the LSM algorithm as follows:

$$\lim_{N\to\infty} \hat{I}_N \le P(S_0; K, T).$$

Since these Greeks are important measures of risk, their efficient computation is needed for a better risk management system. Boyle, Broadie, and Glasserman (1997)

[2]This comes from $p_{t,T} = \mathbb{P}^*[\,S_t^{(i)} \,|\, S_t^{(i)} < Ke^{-r(T-t)}\,] = \mathbb{P}^*[\,\phi < \frac{\ln(K/S_0)-rT+\sigma^2 t/2}{\sigma\sqrt{t}}\,]$ and its adjusting factor $\xi$ is set 1.1 in our simulation study.

suggest three approaches to estimating price sensitivities, including finite-difference approximations, *pathwise* derivatives as direct estimates, and another direct estimates using the *likelihood ratio*. They report that when both direct estimates apply, the pathwise method generally has lower variance. Furthermore, since American options are path-dependent options, our discussion focuses on the application of pathwise derivatives in the case of the American put option.

The pathwise estimate of the true delta $dP/dS_0$ for the American put option is the derivative of the sample price $\hat{P}_i$ with respect to $S_0$. More precisely,

$$\frac{d\hat{P}_i}{dS_0} = \lim_{\varepsilon \to 0} [\hat{P}_i(S_0 + \varepsilon) - \hat{P}_i(S_0)]/\varepsilon,$$

provided that the limit exists with probability 1. If $\hat{P}_i(S_0)$ is computed from the stock price $S_{t_e^{(i)}}$ at the exercise time $t_e^{(i)}$ as follows

$$S_{t_e^{(i)}} = S_0 e^{(r-\sigma^2/2)t_e^{(i)} + \sigma\sqrt{t_e^{(i)}} Z_i}, \qquad Z_i \overset{i.i.d}{\sim} N(0,1), \tag{6}$$

we assume that the stock price $S_{t_e^{(i)}}(\varepsilon)$ is also generated with the same normal shock $Z_i$ as follows

$$S_{t_e^{(i)}}(\varepsilon) = (S_0 + \varepsilon) e^{(r-\sigma^2/2)t_e^{(i)} + \sigma\sqrt{t_e^{(i)}} Z_i}. \tag{6'}$$

As far as the generated stock price $S_{t_e^{(i)}}(\varepsilon)$ is below the optimal exercise boundary stock price and $\hat{P}_i(S_0 + \varepsilon)$ is computed from the stock price $S_{t_e^{(i)}}(\varepsilon)$, we have

$$\frac{d\hat{P}_i}{dS_0} = \frac{d\hat{P}_i}{dS_{t_e^{(i)}}} \frac{dS_{t_e^{(i)}}}{dS_0} = -e^{-rt_e^{(i)}} \mathbf{1}_{\{S_{t_e^{(i)}} < K\}} \frac{S_{t_e^{(i)}}}{S_0}. \tag{7}$$

The next question is whether this estimator is unbiased; that is, whether

$$\mathrm{E}[\, \frac{d\hat{P}_i}{dS_0} \,] = \frac{dP}{dS_0} \equiv \frac{d}{dS_0} \mathrm{E}[\, \hat{P}_i \,]$$

The unbiasedness of the pathwise estimate thus reduces to the interchangeability of derivative and expectation, and the unbiasedness of sample average. Thus the bias of the LSM algorithm implies that of pathwise derivatives.

Applying the same reasoning used above, we obtain the pathwise estimator of vega for the American put option

$$\frac{d\hat{P}_i}{d\sigma} = \frac{d\hat{P}_i}{dS_{t_e^{(i)}}}\frac{dS_{t_e^{(i)}}}{d\sigma} = -e^{-rt_e^{(i)}}\,\mathbf{1}_{\{S_{t_e^{(i)}}<K\}}\left(\ln(S_{t_e^{(i)}}/S_0) - (r-\sigma^2/2)t_e^{(i)}\right). \tag{8}$$

## 5. Simulation Results

In this section, we present the simulation results applying the methods we proposed throughout this paper. Table 1 presents three results for three different initial values of the stock prices, $S_0 = 36, 40, 44$. All of the results are obtained with strike price($K$)=40, volatility($\sigma$)= 0.4, risk-free interest rate($r$)= 0.06, and maturity($T$)= 1 year. The option prices under different regression regimes are presented in columns 4-5, and the simulation results for the *Greeks* are reported in columns 6-9. The rows termed IFDM indicate benchmark option prices and corresponding Greeks. In each panel, we report three different methods to enhance the speed of simulations and three pricing algorithms with the LSM results included for comparison. We mainly examine the case where the option price is in-the-money. The other two cases will be briefly mentioned as they carry similar implications.

Overall, new regressor method overestimates price of options in all situations, while the other two methods underestimate the price. The first panel of Table 1 exhibits the results of different option pricing estimations when the option is in-the-money. Looking at the original method, which conforms to the regressor-selection steps proposed by Longstaff and Schwartz, the pricing errors of the LSM and weighted regression contains less than 1 cent to little above 3 cents. However, the pricing error under the new regressor method looks quite unfavorable. The results of the LSM dominates the other two regressions in this case in terms of relative errors.

However, shifting our interest from the pricing errors to the fitness of the models, we observe a different picture. Figure 9(b) depicts the $R^2s$ of three different regressions. The weighted regression and the new regressor method exhibit a huge improvement in $R^2$ compared to the LSM approach. Especially, the magnitude of difference of $R^2$

between the LSM and the new regressor method is on the order of several multiples. The difference of $R^2$ between the LSM and the weighted regression is also nontrivial. A high $R^2$ indicates that there is a considerable correlation between regressand and regressors, i.e. discounted cash flows and included state variables.

Optimal exercise boundaries for the in-the-money case are presented in Figure 1(c). Two points are worth mentioning. Firstly, near the initial periods, the volatility of optimal exercise boundary increases uncontrollably. Secondly, the gap between the theoretical boundary and the regression-based boundary becomes wider near maturity. These two observations are pathological features of both the LSM approach and our methods. To circumvent these problems, we introduce two corrective procedures, early termination and flexible boundary, as described in Section 3. The simulation results of these procedures are reported in rows 4-9 of Table 1.

Figure 2 illustrates the results when we apply these corrections. Each boundary is computed by the average of 50 estimations. Figure 2(a) is the boundary with no adjustment in either the size of regressors or early termination. As is clear from the figure, both the weighted regression and the new regressor method create steep drops near the early stage of executable steps, while the LSM displays separation from the true boundaries in the middle steps. Figure 2(b) displays the boundary after employing early termination. The flat segment indicates the region where early termination kicks in. From our exercise, we found the excessive volatile segment of the boundary contains not much information on option prices, and skipping estimations on these segments relieves computational burden nontrivially[3] without sacrificing the convergence level, as can be seen from Table 1.

Figures 2(c)-(d) are the results when the flexible boundary is introduced. The flexible boundary affects the number of regressors in different stage of estimation. Figure 8 describes how the number of regressors are varying in different time steps. The cumulative probability under the fixed boundary is always above that of the flexible boundary,

[3]Subfigures (c)-(d) in figures 2-4 show that it reduces the computational burden approximately 30 percentage.

as shown in Figure 8(a). The movement of upper boundaries are depicted in Figure 8(b). In most of the cases, the upper is increasing in time $t$. Compared to the fixed boundary, we have a lower upper bound in early time, but a higher upper bound as time lapses. A higher upper bound tends to enlarge the exercise region and, as a consequence, more observations for estimation will be garnered. Figure 8(c) presents the number of observations in both fixed and flexible boundaries. Clearly, the fixed boundary case stores more observations in early stage than the flexible counterpart, but this situation reverses in later stages[4]. We expect this manipulation can cure the second problem related to the regression-based pricing. Figures 2(c)-(d) do not seem to be a huge success, but progress is undeniable. In particular, the improvement is conspicuous when both the early termination and the flexible boundary are used in the estimation.

Returning to Table 1, the pricing errors under the corrective procedures are at least as good as the original methods and sometimes even better than the original counterparts. Especially, when we apply early termination and flexible boundary restrictions together, the overall pricing errors improve significantly as we expected. The corrective procedures are devised to discard confounding observations, which convey little information on option prices or to smooth the excessive volatile optimal exercise boundary, while the original method does not differentiate them from relevant observations. By implementing these selection procedures, we expect the overall fitness of the models will be improved and pricing errors will become lower. Unlike the improvement shown in pricing errors, the improvement of $R^2$ seems quite limited as shown in Figure 9(a)-(b), but still a significant difference is detected between our proposed regression methods and the LSM counterparts.

The second and third panel of Table 1 and related figures report the results when the option is at-the-money and out-of-the-money respectively. Similar interpretation can be applied as the option is in-the-money.

[4]This phenomena comes from the extension of cumulative probability under the flexible upper bound for regressor's observations through an adjusting factor $\xi$ in step 3 of subsection 3.D.

Figure 5 exhibits the results where all the corrective procedures are applied in three regression frameworks. We would like to emphasize that the weighted regression and new regressor models are tracking the true boundary pretty well, even though there seems to be a noticeable discrepancies from the true boundary in early stage of time.

## 6. Conclusion

In this paper, new algorithms to improve the accuracy of option pricing and computational speed are proposed. Throughout our simulation examples, our results show that the newly proposed algorithm produces a better option pricing than suggested by a simple regression method.

The improvement of our proposed algorithm is two-dimensional: overall fitness of regression and computational speeds. Under the weighted regression and new regressor methods, the $R^2$s increases significantly, and simulation time decreases as less computation should be carried out due to the early termination nature of the algorithm. We believe these features will render a significant difference as state variables used in option pricing regression increase.

Lastly, we do not develop much argument on how to choose adjusting factor, $\xi$. The optimal choice of $\xi$ should be chosen carefully to minimize the gap between the exercise boundary from the theoretical model and that from the flexible boundary model. This is left for future research.

## References

[1] Averbukh, V., 1997, Pricing American options using Monte Carlo simulation, Ph.D. dissertation, Cornell University

[2] Black, F., and M. Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637-659

[3] Bossaerts, P., 1989, Simulation estimators of optimal early exercise, working paper, Carnegie-Mellon University

[4] Boyle, P., M. Broadie, and P. Glasserman, 1997, Monte Carlo methods for security pricing, *Journal of Economic Dynamics and Control* 21, 1267-1321

[5] Broadie, M., and P. Glasserman, 1997, Pricing American-style securities using simulation, *Journal of Economic Dynamics and Control* 21, 1323-1352

[6] Carr, P., 1998, Randomization and the American put, *Review of Financial Studies* 11, 597-626

[7] Cox, J., Ross, S., and Rubinstein, M., 1979, Option Pricing: A Simplified Approach, *Journal of Financial Economics* 7, 220-263

[8] Garcia, D., 1999, A Monte Carlo method for pricing American options, working paper, University of california, Berkeley

[9] Judd, K., 1998, *Numerical Methods in Economics*, MIT Press, Cambridge, Mass.

[10] Longstaff, F. A., and E. S. Schwartz, 2001, Valuing American options by simulation: A simple least-squares approach, *Review of Financial Studies* 14, 113-147

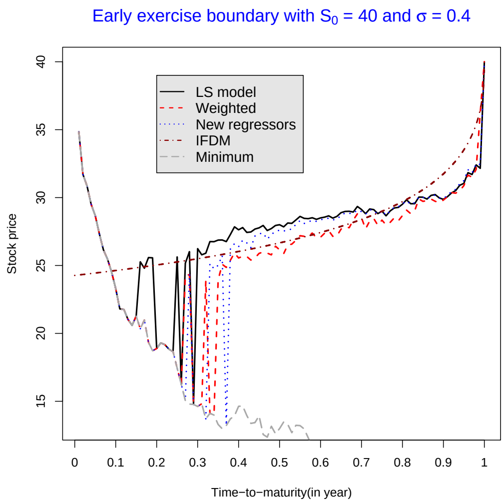
[11] Morokoff, W. J., and R. E. Caflisch, 1995, Quasi-Monte Carlo integration, *Journal of Computational Physics* 122, 218-230

[12] Pizzi, C., and P. Pellizzari, 2002, Monte Carlo pricing of American options - Using nonparametric regression, working paper, University of Venice, Italy

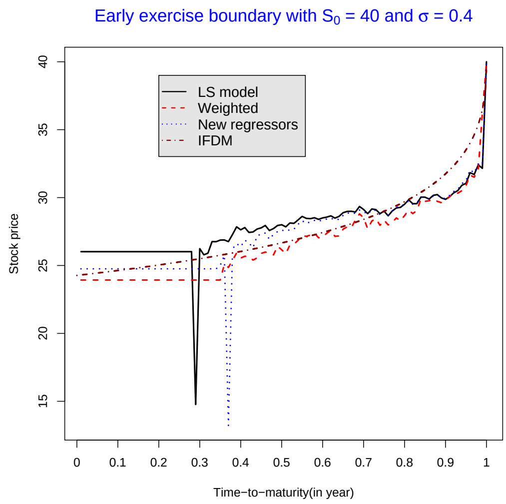
[13] Tilley, J. A., 1993, Valuing American options in a path simulation model, *Transactions of the Society of Actuaries* 45, 83-104

TABLE 1. **Prices and Greeks for the American put option**: The American put option has $K = 40$, $\sigma = 40\%$, $r = .06$, and $T = 1$ and is exercisable 100 times per year. IFDM is estimated with $10^4$ time steps and the same number of underlying state steps. The simulation is based on $2 \cdot 10^4$ ($10^4$ plus $10^4$ antithetic) paths for the underlying price process following the geometric Brownian motion.

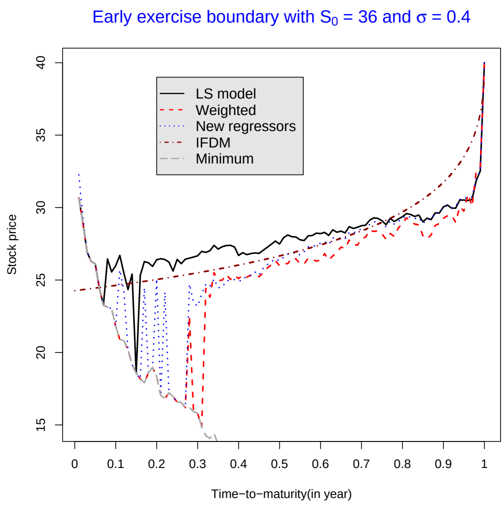
| | Category | Category | Price | | Delta | | Vega | |
|---|---|---|---|---|---|---|---|---|
| $S_0$ | 2 | 1 | Mean | (s.e) | Mean | (s.e) | Mean | (s.e) |
| 36 | IFDM | | 7.108 | | -0.509 | | | |
| | Original | LS | 7.089 | (.018) | -.514 | (.014) | 9.368 | (1.319) |
| | | Weight | 7.077 | (.018) | -.488 | (.014) | 9.214 | (1.232) |
| | | NewReg | 7.351 | (.020) | -.486 | (.015) | 9.630 | (1.225) |
| | Early | LS | 7.088 | (.018) | -.513 | (.014) | 9.364 | (1.319) |
| | Stop | Weight | 7.072 | (.019) | -.487 | (.014) | 9.195 | (1.234) |
| | | NewReg | 7.325 | (.022) | -.484 | (.015) | 9.585 | (1.227) |
| | Stop + | LS | 7.092 | (.017) | -.515 | (.014) | 9.375 | (1.317) |
| | Flexible | Weight | 7.084 | (.019) | -.492 | (.014) | 9.253 | (1.229) |
| | | NewReg | 7.325 | (.019) | -.492 | (.015) | 9.633 | (1.230) |
| 40 | IFDM | | 5.317 | | -0.391 | | | |
| | Original | LS | 5.307 | (.019) | -.394 | (.003) | 10.551 | (0.032) |
| | | Weight | 5.298 | (.017) | -.379 | (.003) | 10.319 | (0.042) |
| | | NewReg | 5.463 | (.020) | -.383 | (.004) | 10.655 | (0.055) |
| | Early | LS | 5.306 | (.020) | -.393 | (.003) | 10.539 | (0.037) |
| | Stop | Weight | 5.295 | (.017) | -.379 | (.003) | 10.304 | (0.046) |
| | | NewReg | 5.453 | (.019) | -.383 | (.004) | 10.635 | (0.061) |
| | Stop + | LS | 5.307 | (.020) | -.395 | (.003) | 10.558 | (0.037) |
| | Flexible | Weight | 5.300 | (.019) | -.381 | (.003) | 10.354 | (0.042) |
| | | NewReg | 5.448 | (.020) | -.386 | (.004) | 10.678 | (0.055) |
| 44 | IFDM | | 3.950 | | -0.296 | | | |
| | Original | LS | 3.944 | (.023) | -.298 | (.003) | 10.979 | (0.050) |
| | | Weight | 3.942 | (.023) | -.289 | (.002) | 10.743 | (0.049) |
| | | NewReg | 4.032 | (.027) | -.295 | (.004) | 11.054 | (0.068) |
| | Early | LS | 3.943 | (.024) | -.298 | (.003) | 10.963 | (0.055) |
| | Stop | Weight | 3.937 | (.024) | -.288 | (.002) | 10.715 | (0.054) |
| | | NewReg | 4.025 | (.026) | -.295 | (.004) | 11.037 | (0.075) |
| | Stop + | LS | 3.945 | (.024) | -.298 | (.003) | 10.984 | (0.055) |
| | Flexible | Weight | 3.941 | (.023) | -.289 | (.002) | 10.758 | (0.053) |
| | | NewReg | 4.022 | (.026) | -.297 | (.004) | 11.077 | (0.070) |



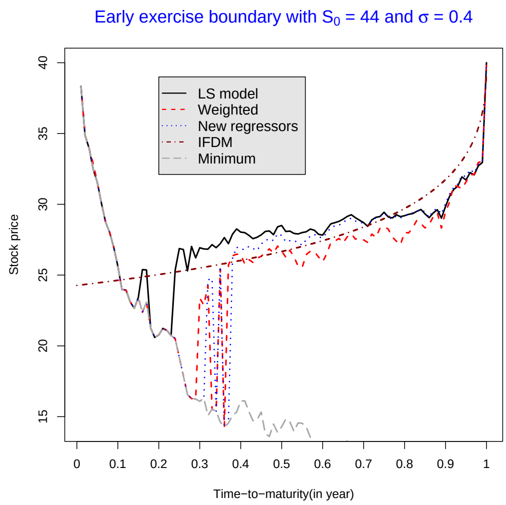

(a) $S_0 = 40$ (at-the-money)




(b) $S_0 = 40$ (at-the-money)

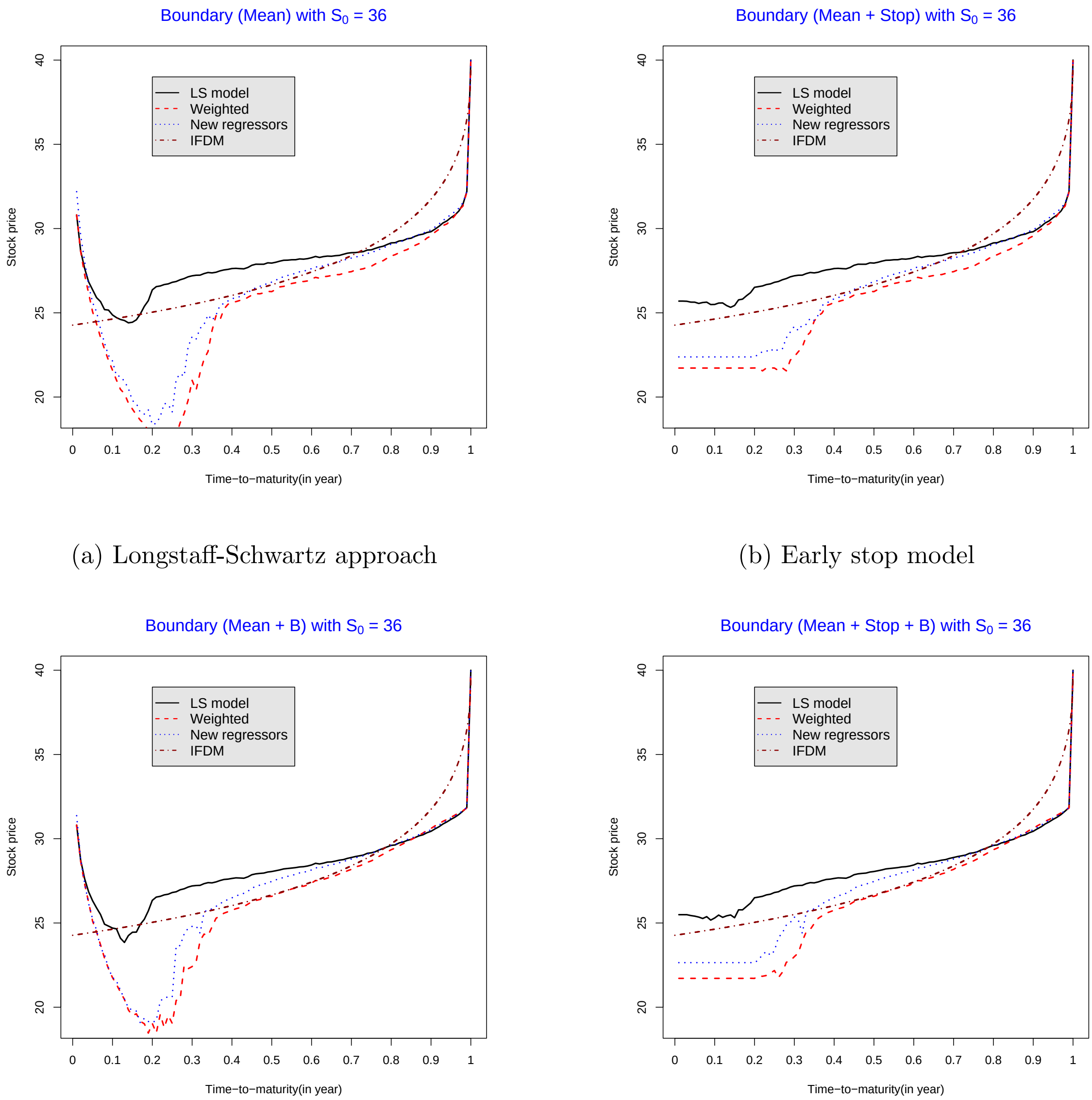



(c) $S_0 = 36$ (in-the-money)




(d) $S_0 = 44$ (out-of-the-money)

FIGURE 1. **Graphs of the optimal exercise boundary for the American put option .** The optimal exercise boundary is shown for the several LSM methodologies w.r.t the time-to-maturity, where the option is exercisable 100 times per year. The benchmark case is obtained by using the implicit FDM (IFDM) with $10^4$ time steps and the same number of underlying state steps. The American put option has $K = 40$, $\sigma = 40\%$, $r = .06$, and $T = 1$. The simulation is based on $2 \cdot 10^4$ ($10^4$ plus $10^4$ antithetic) paths for the underlying price process following the geometric Brownian motion. Minimum is obtained from taking the minimum of the simulated prices at each exercisable time. (b) is obtained from utilizing the decreasing property of optimal exercise boundaries as the time-to-maturity increases.



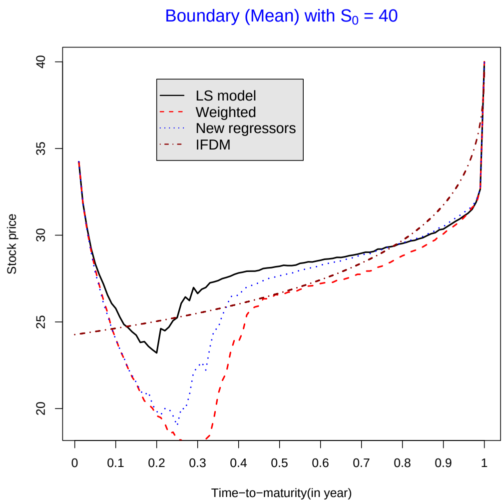

(a) Longstaff-Schwartz approach

(b) Early stop model

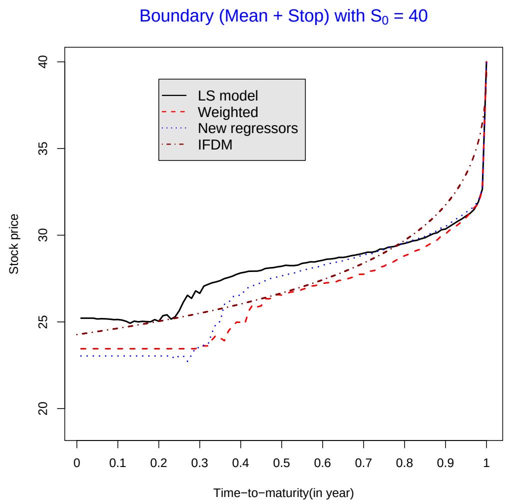
(c) Flexible upper bound

(d) Early stop + Flexible Bound

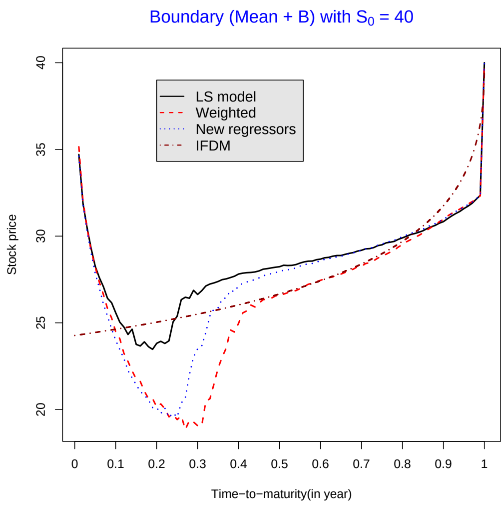
FIGURE 2. **Graphs of the average optimal exercise boundary for the ITM American put option** . The average optimal exercise boundary is shown for the several LSM methodologies, including the original LS approach, early stop model, flexible upper bound model, and the last two-combined model, w.r.t the time-to-maturity, where the option is exercisable 100 times per year. The benchmark case is obtained by using the IFDM with $10^4$ time steps and the same number of underlying state steps. The American put option has $S_0 = 36$, $K = 40$, $\sigma = 40\%$, $r = .06$, and $T = 1$. Each curve is computed from averaging 50 estimated curves, each estimate based on $2 \cdot 10^4$ ($10^4$ plus $10^4$ antithetic) paths for the underlying price process. Each different model uses the same simulated data.



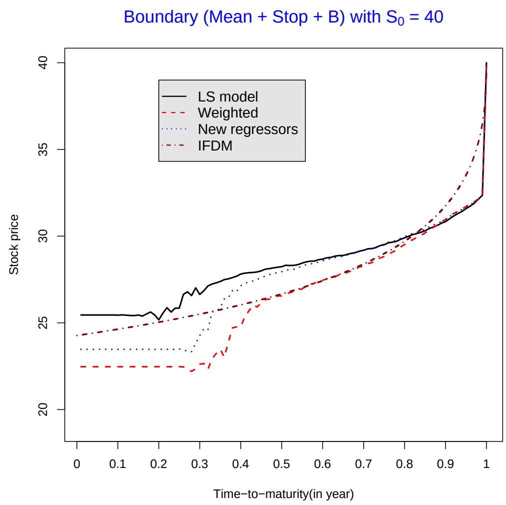

(a) Longstaff-Schwartz approach




(b) Early stop model

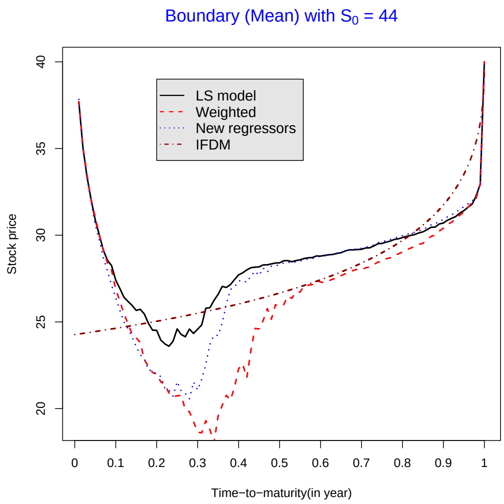



(c) Flexible upper bound

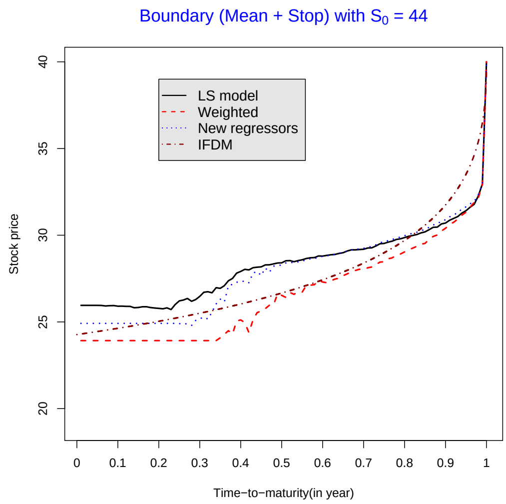



(d) Early stop + Flexible Bound

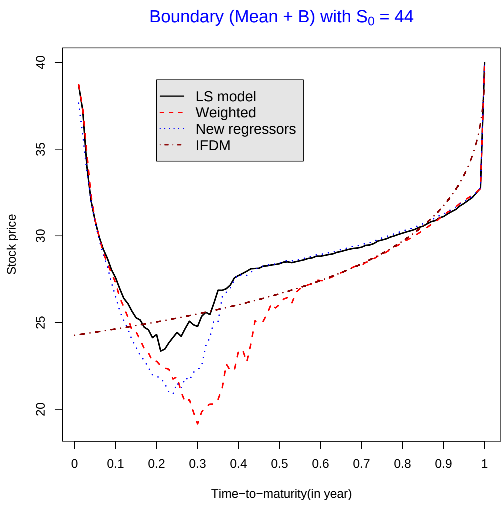
FIGURE 3. **Graphs of the average optimal exercise boundary for the ATM American put option** . The average optimal exercise boundary is shown for the several LSM methodologies, including the original LS approach, early stop model, flexible upper bound model, and the last two-combined model, w.r.t the time-to-maturity, where the option is exercisable 100 times per year. The benchmark case is obtained by using the IFDM with $10^4$ time steps and the same number of underlying state steps. The American put option has $S_0 = 40$, $K = 40$, $\sigma = 40\%$, $r = .06$, and $T = 1$. Each curve is computed from averaging 50 estimated curves, each estimate based on $2 \cdot 10^4$ ($10^4$ plus $10^4$ antithetic) paths for the underlying price process. Each different model uses the same simulated data.



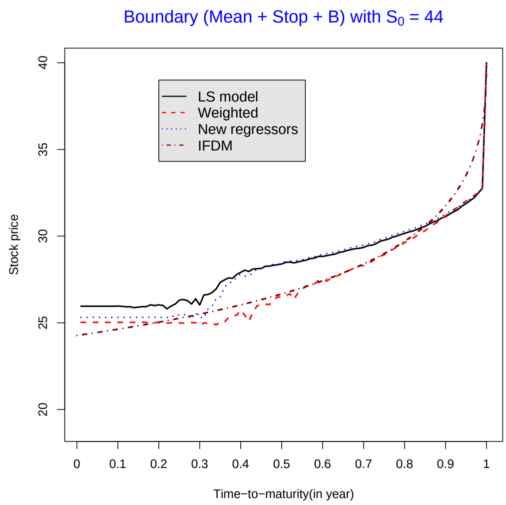

(a) Longstaff-Schwartz approach




(b) Early stop model

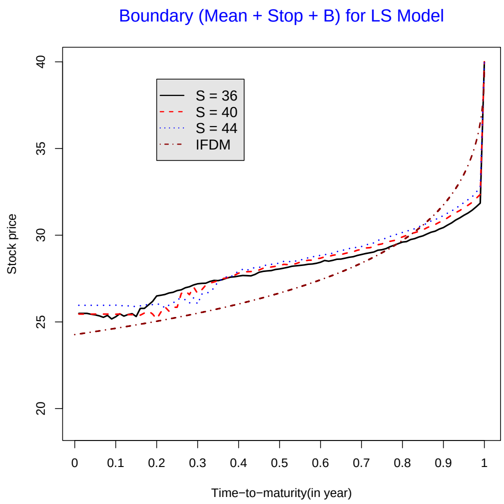



(c) Flexible upper bound

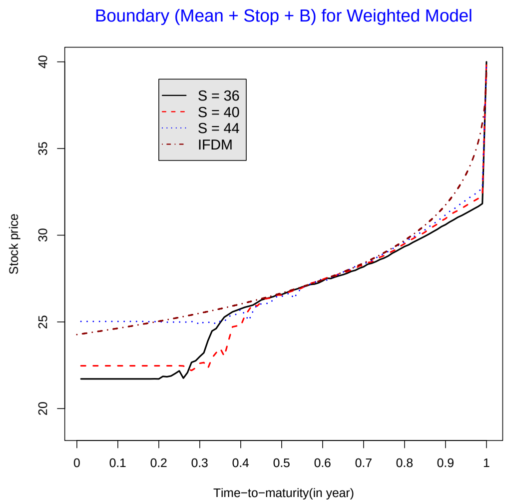



(d) Early stop + Flexible Bound

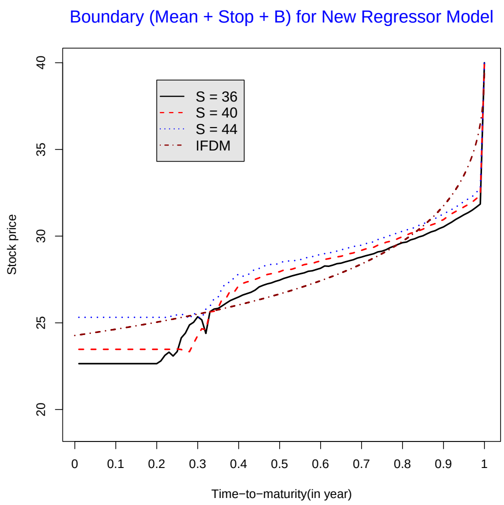
FIGURE 4. **Graphs of the average optimal exercise boundary for the OTM American put option .** The average optimal exercise boundary is shown for the several LSM methodologies, including the original LS approach, early stop model, flexible upper bound model, and the last two-combined model, w.r.t the time-to-maturity, where the option is exercisable 100 times per year. The benchmark case is obtained by using the IFDM with $10^4$ time steps and the same number of underlying state steps. The American put option has $S_0 = 44$, $K = 40$, $\sigma = 40\%$, $r = .06$, and $T = 1$. Each curve is computed from averaging 50 estimated curves, each estimate based on $2 \cdot 10^4$ ($10^4$ plus $10^4$ antithetic) paths for the underlying price process. Each different model uses the same simulated data.




(a) Longstaff-Schwartz Model

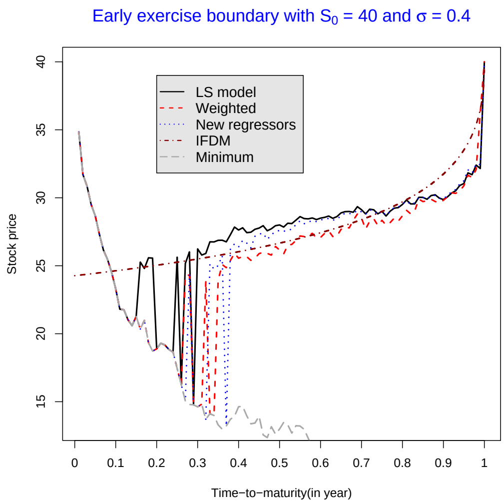



(b) Weighted Model

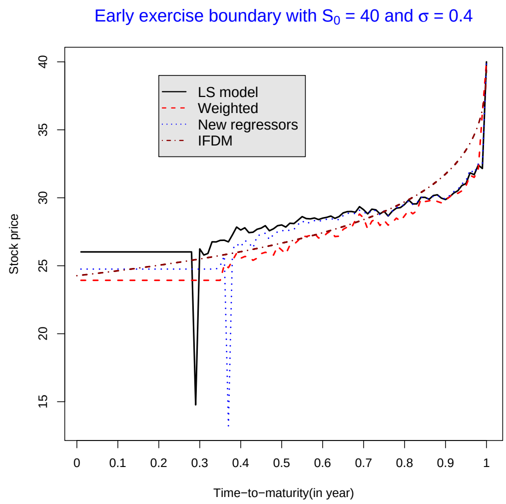



(c) New Regressor Model

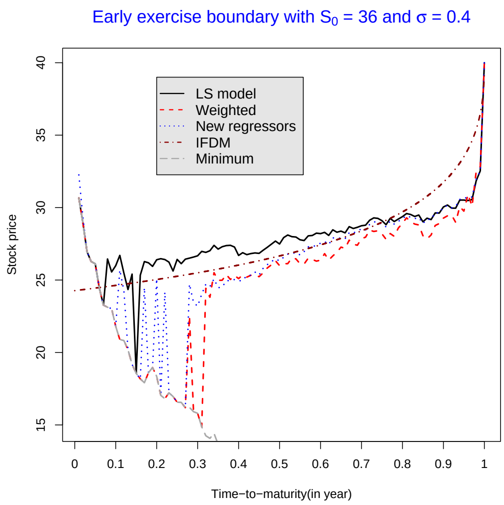
FIGURE 5. **Graphs of the average optimal exercise boundary for the American put option under the combined model of early stop and flexible boundary**. The average optimal exercise boundary is shown w.r.t the time-to-maturity, where the option is exercisable 100 times per year. The benchmark case is obtained by using the IFDM with $10^4$ time steps and the same number of underlying state steps. The American put option has $S_0 = 36, 40, 44$, $K = 40$, $\sigma = 40\%$, $r = .06$, and $T = 1$. Each curve is computed from averaging 50 estimated curves, each estimate based on $2 \cdot 10^4$ ($10^4$ plus $10^4$ antithetic) paths for the underlying price process. Each different model uses the same simulated data.



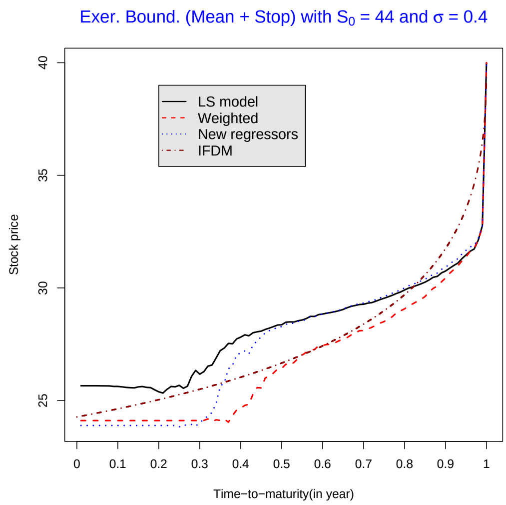

(a) $S_0 = 40$ (at-the-money)




(b) $S_0 = 40$ (at-the-money)

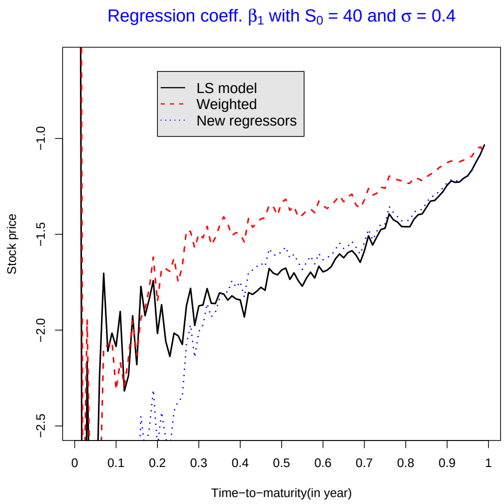



(c) $S_0 = 36$ (in-the-money)

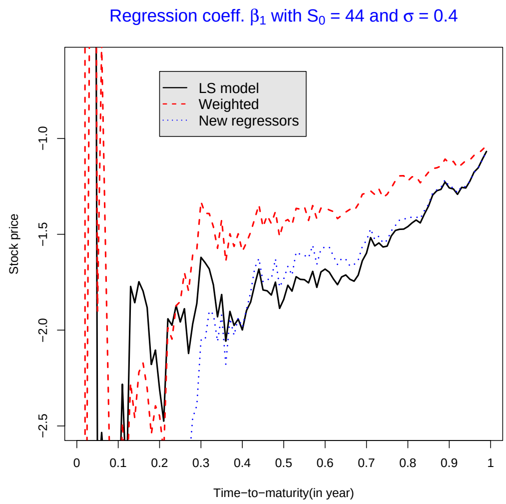



(d) $S_0 = 44$ (out-of-the-money)

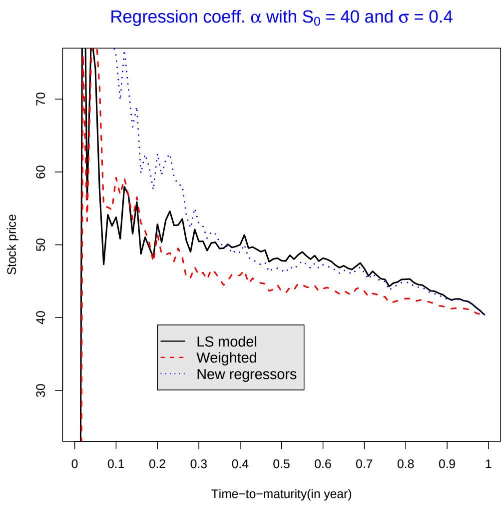
FIGURE 6. **Graphs of the average optimal exercise boundary for the American put option** . The average optimal exercise boundary is shown for the several LSM methodologies w.r.t the time-to-maturity, where the option is exercisable 100 times per year. The benchmark case is obtained by using the implicit FDM (IFDM) with $10^4$ time steps and the same number of underlying state steps. The American put option has $K = 40$, $\sigma = 40\%$, $r = .06$, and $T = 1$. Each curve is computed from averaging 100 estimated prices at each exercisable time, each estimate based on $2 \cdot 10^4$ ($10^4$ plus $10^4$ antithetic) paths for the underlying price process. Minimum is obtained from taking the minimum of the simulated prices at each exercisable time. (b) is obtained from utilizing the decreasing property of optimal exercise boundaries as the time-to-maturity increases.



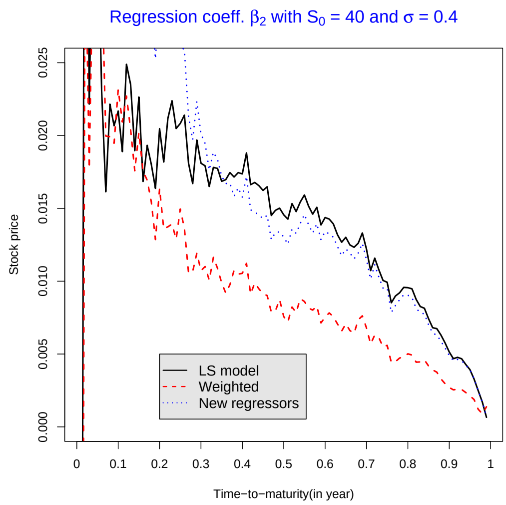

(a) $\beta_1$ with $S_0 = 40$




(b) $\beta_1$ with $S_0 = 44$

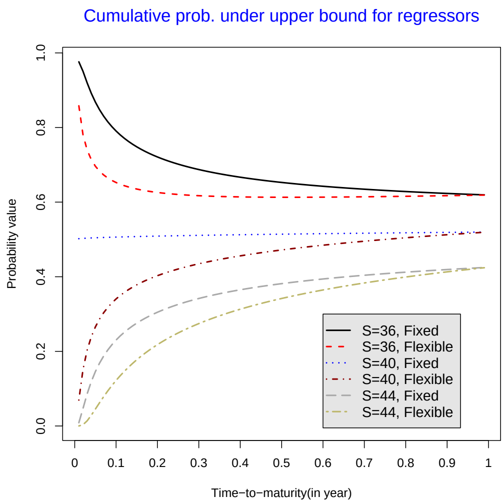



(c) $\alpha$ with $S_0 = 40$

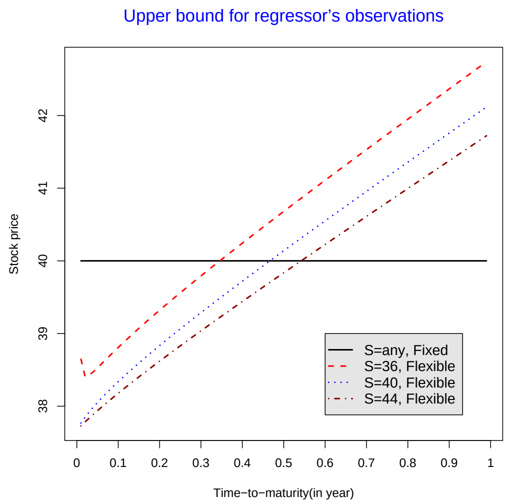



(d) $\beta_1$ with $S_0 = 40$

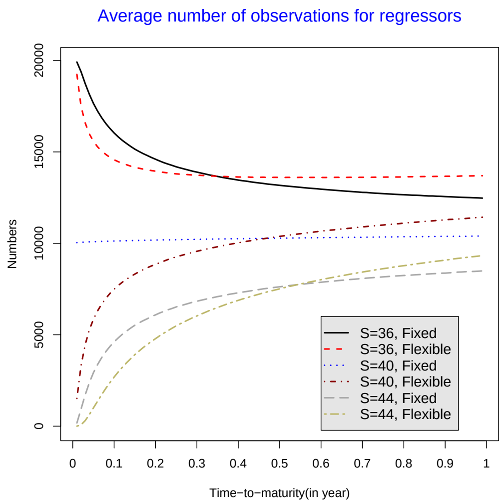
FIGURE 7. **Graphs of the coefficient for the regression** . The coefficient for the regression optimal exercise boundary is shown for the several LSM methodologies w.r.t the time-to-maturity, where the option is exercisable 100 times per year. The benchmark case is obtained by using the implicit FDM (IFDM) with $10^4$ time steps and the same number of underlying state steps. In this comparison, the strike price of the put is 40, the short-term interest rate is .06, the underlying stock price $S_0$, the volatility of returns $\sigma = 40\%$, and the year until the final expiration $T = 1$. The simulation is based on $2 \cdot 10^4$ (100,000 plus 100,000 antithetic) paths for the underlying price process following the geometric Brownian motion. Minimum is obtained from taking the minimum of the simulated prices at each exercisable time. (b) is obtained from utilizing the decreasing property of optimal exercise boundaries as the time-to-maturity increases.




(a) Cumulative probability under pseudo upper bound

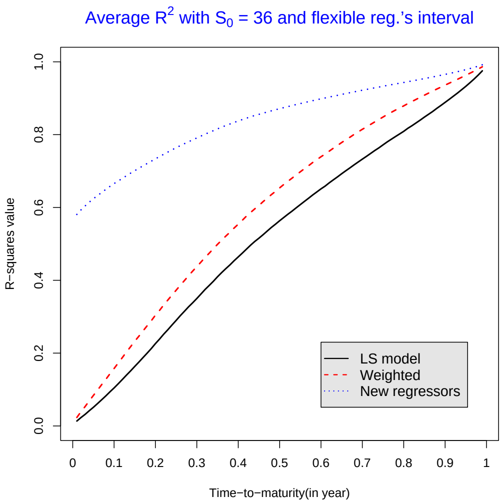



(b) Upper bound for regressor's observations

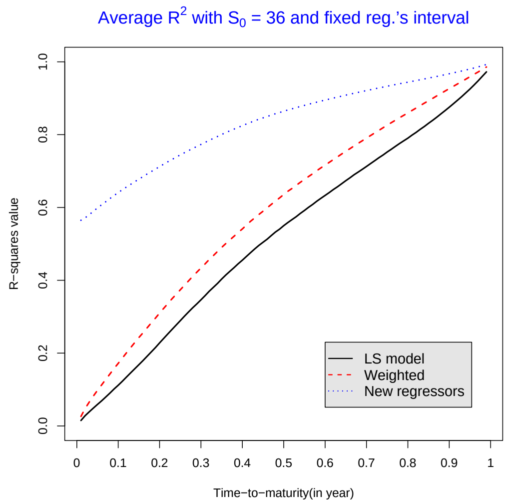



(c) # of observations for regressors

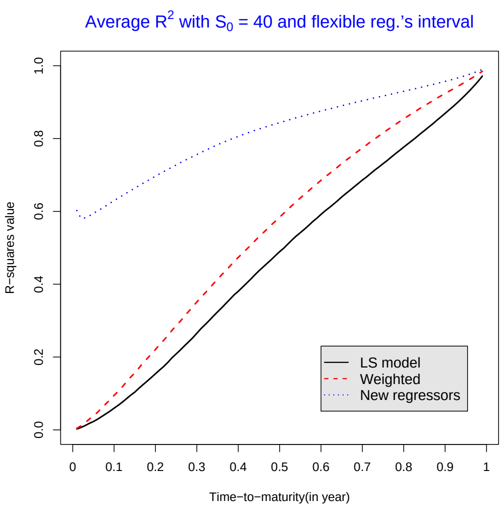
FIGURE 8. **Graphs of summary statistics for independent regressors**. (a) describes the cumulative probability, $p_{t,T} = \mathbb{P}^*[\, S_t^{(i)} \,|\, S_t^{(i)} < Ke^{-r(T-t)} \,]$ at each decision time $t$ for simulated paths. (b) describes the upper bound for regressor's observations in the linear regression model at each decision time $t$, which is $U_r(t) = K = 40$ for the fixed case and $U_r(t) = S_0 e^{(r-\sigma^2/2)t+\sigma\sqrt{t}q_{t,T}}$ for the flexible case, where $q_{t,T}$ satisfies $\mathbb{P}^*[\, \phi \leq q_{t,T} \,] = 1.1 p_{t,T}$ with $\phi \sim N(0,1)$. (c) describes the observation numbers at the decision time $t$ satisfying $S_t^{(i)} < U_r(t)$ with the initial stock price $S_0 = 36, 40, 44$.



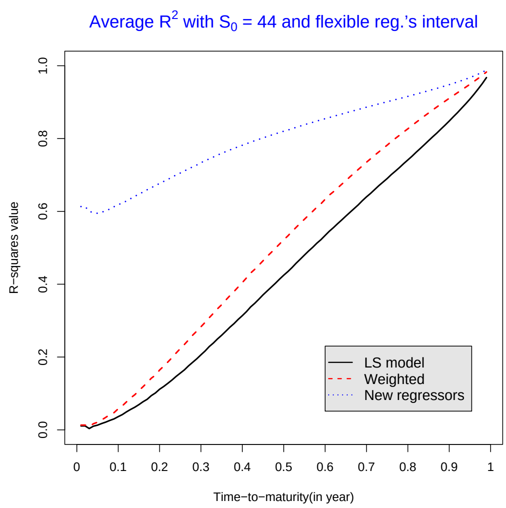

(a) $R^2$ with $S_0 = 36$ and flexible




(b) $R^2$ with $S_0 = 36$ and fixed




(c) $R^2$ with $S_0 = 40$ and flexible




(d) $R^2$ with $S_0 = 44$ and flexible

FIGURE 9. **Graphs of average $R^2$s for regressions**. The average $R^2$ values for the regression are shown for the several LSM methodologies, including the original LS approach, weighted regression model, and new regressor model and each model with (fixed or flexible) upper bound for the independent regressor variable with respect to the time-to-maturity, where the option is exercisable 100 times per year. The American put option has $K = 40$, $\sigma = 40\%$, $r = .06$, and $T = 1$. Each curve is computed from averaging 50 estimated $R^2$ values at each exercisable time $t_e$, each estimate computed from the linear regression model with the independent regressors, where numbers of their observation are dependent on the fixed $(K)$ or flexible $(S_0 e^{(r-\sigma^2/2)t_e + \sigma\sqrt{t_e} q_{t_e,T}})$ upper bound.