Valuing American Derivatives by Least Squares Methods

Mario Cerrato*

September 8, 2009

*Corresponding author. University of Glasgow, Department of Economics, email address: m.cerrato@lbss.gla.ac.uk. The autor wishes to thank Carol Alexander, Gabriel Talmain and Guglielmo Maria Caporale for valuable comment. I am also grateful to Kan Kwok Cheung for carrying preliminary computations that helped me form some of the ideas analysed in this paper. The usual disclaimer applies.

Abstract

Least Squares estimators are notoriously known to generate sub-optimal exercise decisions when determining the optimal stopping time. The consequence is that the price of the option is underestimated. We show how variance reduction methods can be implemented to obtain more accurate option prices. We also extend the Longstaff and Schwartz (2001) method to price American options under stochastic volatility. These are two important contributions that are particularly relevant for practitioners. Finally, we extend the Glasserman and Yu (2004b) methodology to price Asian options and basket options.

Key Words: American options, Monte Carlo method
JEL Classification: G10, G13

I Introduction

Monte Carlo method to price American options seems to be an active research area. The reason is mainly due to its flexibility to price path dependent options and to solve high dimensional problems.

It is now standard combining Monte Carlo methods and regression methods to price derivatives with American features. For example, Longstaff and Schwartz (2001) suggest using Least Squares approximation to approximate the option price on the continuation region and Monte Carlo methods to compute the option value (LS). Proofs of convergence of Monte Carlo estimators are derived under various assumptions. Therefore a small sample analysis is necessary. For example the proof in Longstaff and Schwartz (2001) paper is based on various general assumptions.

Clement et al (2002) show that the option price converges, in the limit, to the true price. However the theoretical proof in Clement et al (2002) has also some limitations since it is based on a sequential rather than joint limit.

Glasserman et al (2004a) consider the limitations in Clement et al (2002) and prove convergence of the LS estimator as the number of paths and the number of polynomials functions increase together. A further assumption of martingales polynomials is required in this case.

Glasserman et al (2004b) (GY) implement a weighted Monte Carlo Estimator (WME) to price American derivatives and show that their estimator produces less disperse estimates of the option price. However, no finite-sample proof of convergence of the proposed estimator is provided in that study. Furthermore, the proof of Theorem 1 is based on a two period framework.

Applications of Monte Carlo estimators to price financial derivatives generally require using variance reduction techniques. One common feature to some of the studies cited above is that they only consider antithetic variates. As we shall see, particularly when pricing American style derivatives using one method rather than another makes the difference when determining the early exercise value.

In this paper we analyze the finite sample approximation of the LS (2001) and GY (2004b) estimators by extending empirical studies such as Stentoft (2004)[1].

[1]Note that, although we also evaluate competitive models (i.e. Longstaff-Schwartz, 2001 and

As shown in Glasserman and Yu (2004a) the choice of the basis function used in the regression is very important since (uniform) convergence of the option price to the true price can only be guaranteed if polynomials span the "true optimum". To address this issue, we consider different basis functions and suggest possible "optimal polynomials" for specific applications. We then discuss ways to implement variance reduction techniques in this context and study the contribution of these methods to variance reduction and bias[2]. Finally, this study proposes a very simple and flexible approach, that extends the LS Monte Carlo estimator, to price American options under stochastic volatility. This is an important novel contribution. We show that our method produces very accurate option prices.

II The Least Squares Monte Carlo Methods

We consider a probability space (Ω, F, P) and its discrete filtration $(F_i)_{i=0,\ldots,n}$, with n being an integer representing the number of time points considered. Define with $X_0, X_1, ..., X_n$ an R^d valued Markov chain representing the state variable recording all the relevant information on the price of an underlying asset. Assume that $V_i(x)$, $x \in R^d$, is the value of an option if exercised at time i under the state x. Using a dynamic programming framework the value of the option is given by:

$$V_i(x) = \sup_{\tau \in \Gamma} E[\Theta_\tau(X_\tau)|X_i = x] \tag{1}$$

$$V_n(x) = \Theta(x) \tag{2}$$

with

$$V_i(x) = \max\{\Theta_i(x), E[V_{i+1}(X_{i+1})|X_i = x]\} \tag{3}$$

Glasserman and Yu, 2004b), this is not the primary objective of the paper. In fact, any sort of comparison to make statistical (economic) sense should report a very large number of results that span the space of models that need Monte Carlo simulations. Clearly this is not our primary goal. Instead we put emphasis on extending variance reductions techniques. In fact, although some of the methodologies presented in this paper have also been considered in Rasmussen (2005) and Broadie and Cao (2008), we believe there is scope for further investigation.

[2]Abadir and Paruolo (2008) propose original modifications of anthitetic variates and control variates for dynamic models.

where the expectation above is taken under the risk neutral measure.

At expiry date if the option is exercised the option holder will receive the payoff $\Theta(X)$ (see Equation 2). Prior the expiry date the option value is given by the maximum between the payoff provided by the option if immediately exercised and the continuation value (see Equation 3). As it stands Equation (3) is of little use as it is not directly applicable. In fact, the conditional expectation is hard to compute in this specific case. However, if we assume that the option value is a square integrable function, then $V_i(.)$ will be a function spanning the Hilbert space and one can approximate the conditional expectation in (3) by the orthogonal projection on the space generated by a finite number of basis functions ϕ_{ik}, $i = 1, 2, ..., n$ and $k = 0, 1, ..., K$, such that

$$V_n = \phi_n(x) \tag{4}$$

$$V_i(x) = \max\{\phi_i(x), E[(V_{i+1}(X_{i+1})|X_i = x]\} \tag{5}$$

One can now compute the conditional expectation by a simple regression approach:

$$V_{i+1}(X_{i+1}) \equiv \sum_{k=0}^{K} c_{ik}\phi_{ik}(X_i) + \varepsilon_{i+1} \tag{6}$$

Thus, we replace the difficult problem of solving (1)-(3), with a simple regression requiring the estimation of $K + 1$ coefficients in (6).

Definition 1 *In Equation 6 we have included the error term* ε_i *. As pointed out in Grasserman and Yu (2004b), this approximation will be exact if Definition 2 below holds.*

Definition 2 *If* $E(\varepsilon_{i+1}|X_i) = 0$ *and* $E[\phi_i(X_i)\phi_i(X_i)']$ *is non-singular, then* $V_i^* \longrightarrow V_i$ *for all* $i = 0, 1, ..., n$ *, where* V_i^* *is the estimated option price using the LS regression above.*

Proof of convergence of this estimator is provided in LS (2001), but it applies to the simplest possible case of only one exercise time and one state variable. Clement et al (2002) consider a multi period framework under the assumption that the number of basis functions is fixed. This means that the regression used is correct, therefore no sample bias is considered. GY (2004a) generalize the proof in Clement et al (2002) and show that the option price, using regression methods, converges to the true price as the number of basis functions and the number of Monte Carlo replications (M), $(K, M)_{\longrightarrow\infty}$. But martingales basis must be used in this case.

All the theoretical results mentioned above are very important, particularly from a theoretical point of view. However, for practical applications of these methodologies we are more concerned with their performance in finite sample.

In Equation (6) the conditional expectation is approximated using current basis functions (that is $\phi_i(X_i)$). However one would expect the option price at time $i+1$ to be more closely correlated with the basis function $\phi_{i+1}(X_{i+1})$ rather than $\phi_i(X_i)$. GY (2004b) develop a method based on Monte Carlo simulations where the conditional expectation is approximated by $\phi_{i+1}(X_{i+1})$ rather than $\phi_i(X_i)$. They show that their Monte Carlo scheme has a regression representation given by:

$$V_{i+1}^{*}(X_{i+1}) \equiv \sum_{k=0}^{K} \varpi_{ik}\phi_{i+1,k}(X_{i+1}) + \varepsilon_{i+1} \tag{7}$$

The application of this methodology requires defining martingale basis, $E(\phi_{i+1}(X_{i+1})|X_i) = \phi_i(X_i)$ for all i. GY (2004b) call this method regression later, since it involves using functions $\phi_{i+1}(X_{i+1})$. On the other hand, they call the LS (2001) method regression now since it uses functions $\phi_i(X_i)$. Although Theorem 1 in GY (2004b) provides a justification for using regression later as opposed to regression now, its proof is based on a single period framework. Furthermore, GY (2004b) neither provide an empirical application nor suggest ways of obtaining martingale basis.

To start discussing one of the objectives of this paper (i.e. using variance reduction techniques), we start with a simple example and estimate early exercises values for American put options by crude Monte Carlo methods and using the control variate method described in section VI.[3] Table 1 shows the empirical results

We have used Equations (4)-(6) to compute the price of the option and considered three in-the-money put options with strike $45, initial price $40, maturity seven

[3] We use the Longstaff and Schwartz (2001) method.

Monte Carlo	EU-BS	ExerciseM	Binomial	ExerciseB	Difference
5.265	4.84	0.425	5.265	0.425	0.00%
6.234	5.96	0.274	6.244	0.284	-3.50%
7.374	7.14	0.234	7.383	0.243	-3.80%
Antithetic V.					
5.261	4.84	0.421	5.265	0.425	-0.95%
6.241	5.96	0.281	6.244	0.284	-1.40%
7.384	7.14	0.244	7.383	0.243	-0.41%
Control V.					
5.264	4.84	0.424	5.265	0.425	-0.23%
6.246	5.96	0.286	6.244	0.284	0.70%
7.387	7.14	0.247	7.383	0.243	1.62%

Monte Carlo refers to the crude Monte Carlo method. EU-BS is the price of an equivalent European option using Black and Scholes. Binomial refers to the methodology used. ExerciseM and ExerciseB are respectively early exercises prices using Monte Carlo and Binomial methods.

Table 1: Monte Carlo

months, risk free rate of interest 4.88%p.a. and volatilities 20%, 30% and 40% respectively. The last column shows the difference, in percentages, between estimates of the early exercise value by crude Monte Carlo, Monte Carlo with control variate, and Binomial methods. Variance reduction techniques reduce the bias by an order of 1.6% on average. This is likely to have a substantial impact on the estimate of the put option price and consequently on the price of large books of options.

III Valuing American Put Options

Thus, it is important to implement Monte Carlo methods using variance reduction techniques since it reduces the bias in the estimation of the early exercise value and we can compute a more accurate option price. Variance reduction techniques may help to reduce the probability of generating sub-optimal exercise decisions. In this section we first apply the LS (2001) and GY (2004b) methods to price American put options and then implement the same methodologies using different basis functions and different variance reduction techniques. As pointed out the choice of the basis functions is very important since (uniform) convergence can only be guaranteed if the polynomial used is an "optimal polynomial".

We start with a simple application which is in line with previous works (see for example Longstaff and Schwartz, 2001 and Stentoft, 2004) where we use standard antithetic variance. Prices reported are averages of 50 trials. We report standard errors and root mean square errors as a measure of the small sample bias and bias in the estimation of the conditional expectation in (6). As a benchmark, we consider a Binomial method with 10,000 time steps. Table 2 shows the empirical results. To implement the GY (2004b) estimator we specify the following martingale basis under geometric Brownian motion and exponential polynomial:

$$\phi_{ik}(X_i) = (X_i)^k \exp -(kr + k(k-1)\sigma^{2/2})(t_i - t_0) \tag{8}$$

On the other hand we could not find a valid martingale specification for polynomials when Laguerre basis were used. Finally, following GY (2004a) Hermite polynomials (H_k) are martingales after the transformation in (9)

$$\phi_{ik}(X_i) = t^{k/2} H_k \tag{9}$$

		Exponential			Laguerre			Binomial
Basis Functions		2	3	4	2	3	4	
Methodology	Vol./Mat.							
GY	0.3/0.333	5.694	5.704	5.709	0	0	0	5.706
SE		[0.0106]	[0.0098]	[0.0131]				
RMSE		[0.7172]	[0.7300]	[0.7309]				
LS	0.3/0.333	5.694	5.6991	5.703	5.689	5.699	5.706	
SE		[0.0092]	[0.0115]	[0.0131]	[0.01321]	[0.01385]	[0.0012]	
RMSE		[0.7185]	[0.7279]	[0.7338]	[0.01735]	[0.0061]	[0.0047]	
GY	0.3/0.583	6.223	6.238	6.246	0	0	0	6.244
SE		[0.0143]	[0.0195]	[0.0149]				
RMSE		[1.2689]	[1.2838]	[1.2875]				
LS	0.3/0.583	6.221	6.231	6.244	6.227	6.243	6.234	
SE		[0.0068]	[0.0155]	[0.0133]	[0.0118]	[0.0133]	[0.0143]	
RMSE		[1.274]	[1.2866]	[1.2818]	[0.01678]	[0.0013]	[0.0059]	
GY	0.4/0.333	6.491	6.511	6.513	0	0	0	6.51
SE		[0.0152]	[0.0166]	[0.0123]				
RMSE		[1.521]	[1.537]	[1.5386]				
LS	0.4/0.333	6.499	6.501	6.512	6.495	6.511	6.514	
SE		[0.0073]	[0.0141]	[0.01898]	[0.0152]	[0.0148]	[0.0172]	
RMSE		[1.522]	[1.532]	[1.5412]	[0.0146]	[0.0010]	[0.0037]	
GY	0.4/0.583	7.363	7.378	7.382	0	9	0	7.383
SE		[0.0191]	[0.0254]	[0.0238]				
RMSE		[1.4086]	[1.4287]	[1.4321]				
LS	0.4/0.583	7.37	7.376	7.382	7.362	7.376	7.374	
SE		[0.0078]	[0.0217]	[0.0178]	[0.0248]	[0.0013]	[0.0129]	
RMSE		[1.4107]	[1.421]	[1.4301]	[0.0209]	[0.5611]	[0.0095]	

Table 2: Longstaff-Schwartz (2001), Glasserman-Yu (2004b) Methods

Note that GY and LS are respectively the Longstaff-Schwartz and Glasserman-Yu methodologies. Zeros refer to cases when we were not able to implement the GY method. SE and RMSE are respectively standard errors and root mean square errors.

We use in this case Equation (7) instead of Equation (6) to estimate the conditional expectation. The first column in the Table 2 shows the methodologies used (i.e. Glasserman and Yu, 2004b and Longstaff and Schwartz, 2001). The second column shows volatilities and time to expiry of the options. The strike is assumed to be \$45 and the initial stock price \$40. Therefore we only consider in the money options. The risk free rate of interest is assumed to be 4.88% p.a. Fifty time steps are considered in combination with 100,000 Monte Carlo replications. We use two different polynomial basis, namely exponential and Laguerre of order two, three and four. Following Brodie and Kaya (2004) the RMSE is defined as[4] $(bias^2 + variance)^{1/2}$. The results in the Table 2 suggest that, in general, Laguerre polynomials are appro-

[4] Refer to Broadie and Kaya (2004) for details

priate. Small RMSEs and standard errors support this conclusion. RMSEs confirm what has been found in other empirical studies. That is, the convergence of these estimators is not uniform, and increasing the number of basis does not necessarily reduce the bias.

IV Regression Methods and Moment Matching

One important issue when pricing derivatives by simulation is that we can confidently price the option if, in the first place, we have correctly simulated the dynamics of the underlying asset. Moment matching helps achieving this goal. This particular variance reduction technique has never been considered in the type of applications as the ones in this paper. Therefore, it is of some interest, especially for practitioners, to see if it is suitable for these applications.

We follow Boyle et al (1997) and consider an R^d valued Markov chain sequence of simulated paths $X_0, X_1, ..., X_M$ (with M being the number of Monte Carlo simulations) and assume that we know the expectation $E(X) = \exp(-rt)X_0$. The sample mean process of the sequence can be written as:

$$\overline{X(t)} = \frac{1}{M}\sum_{j=1}^{M} X_j \tag{10}$$

In finite sample we know that $E[X(t)] \neq \overline{X}(t)$. However we can adjust the simulated paths such that the sample mean is equal to $E[X(t)]$

$$\widetilde{X}_j(t) = X_j(t) + E[X(t)] - \overline{X}(t) \tag{11}$$

where $\widetilde{X}$is the new simulated path after the transformation.

Consequently, we have that $E[(\widetilde{X}(t)] = E[(X(t)]$ and the mean of the simulated sample path matches the population mean exactly. Apart from matching the first moment of the process, we can also match higher order moments such as the variance for example. In this case we can re-write the process in (11) as

$$\widetilde{X}_j(t) = [X_j(t) - \overline{X}(t)]\frac{\sigma_X}{s_X} + E[X(t)] \tag{12}$$

where σ_X and s_X are, respectively, the population and the sample variance.

One important drawback of the process in (12) is that sample paths are correlated and therefore it is unlikely that the initial and the simulated processes will have the same distribution. The correlation also makes estimates of standard errors meaningless. To overcome these drawbacks, in the empirical application, we extend this variance reduction technique as showed in Glasserman (2004) and implement the additive process in (12) to the standard Brownian motion process $W(t)$, in the following way:

$$\widetilde{W}_j(t) = [W_j(t) - \overline{W}(t)]/\frac{s_W}{\sqrt{t}}$$

where $\overline{W}(t)$ is the mean of W, and s_W is the standard error.

To preserve independence between sample paths, we also re-scale the increments of the process $W(t)$ as follows:

$$\widetilde{W}_j(t_i) = \sum_{i=1}^{n} \sqrt{t_i - t_{i-1}}\frac{Z_{ij} - \overline{Z}_j}{s_j}$$

where Z_{ij} are standard normal variables, $\overline{Z}_j$ is the mean of Z, and $s_j^2 = \frac{1}{M-1}\sum_{j=1}^{M}(Z_{ij} - \overline{Z_j})^2$.

V Empirical Results

We only consider one of the options presented in Table 2[5]. We consider a put option with seven months to expiry, volatility 40%, initial stock price \$40. The rate of interest is 4.88%p.a. We set the number of steps equal to 50 in all the experiments and compute standard errors and root mean squares errors for sample size of 16, 70, 300, 1000 based on 2000 simulations. Values are reported in log term.

[5] More cases were analysed but not reported to save space. Results are more or less unaffected. These simulations are available upon request.

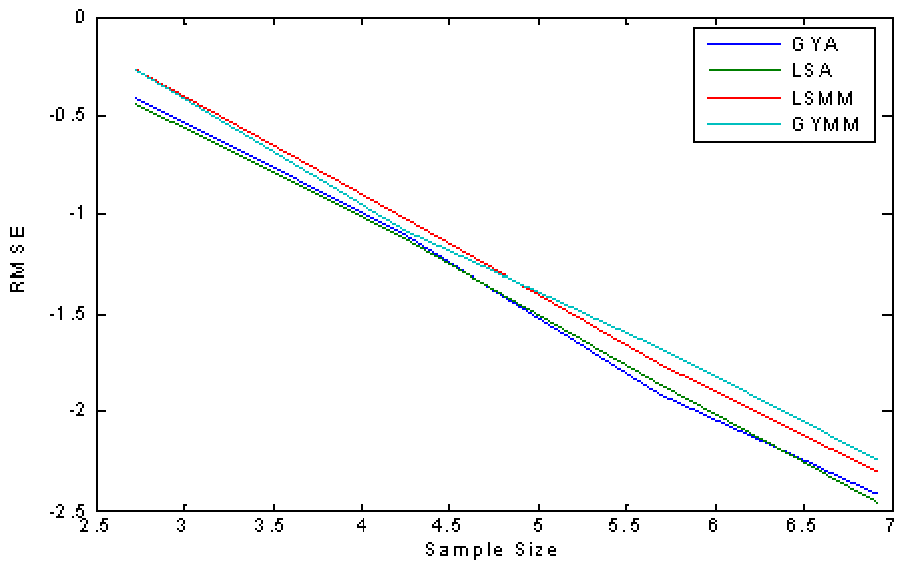

Figure 1: Standard errors versus sample size in pricing an American Put option with strike $45 and initial stock price $40.

In Figure 1 we compare standard errors versus sample size for GY (2004b) and LS (2001) methods using antithetic variates (A) and moment matching (MM). Antithetic variates outperform moment matching in this case. This becomes particularly relevant as the sample increases. Interestingly standard errors for GY (2004b) and LS (2004) methods are narrower when moment matching is used (i.e. almost indistinguishable).

Root mean squares errors versus sample size are reported in Figure 2. Antithetic variates outperforms moment matching as the sample size increases. Thus, in this case, standard antithetic variates seem to do better than moment matching. In the next section we shall present an alternative approach based on control variates.

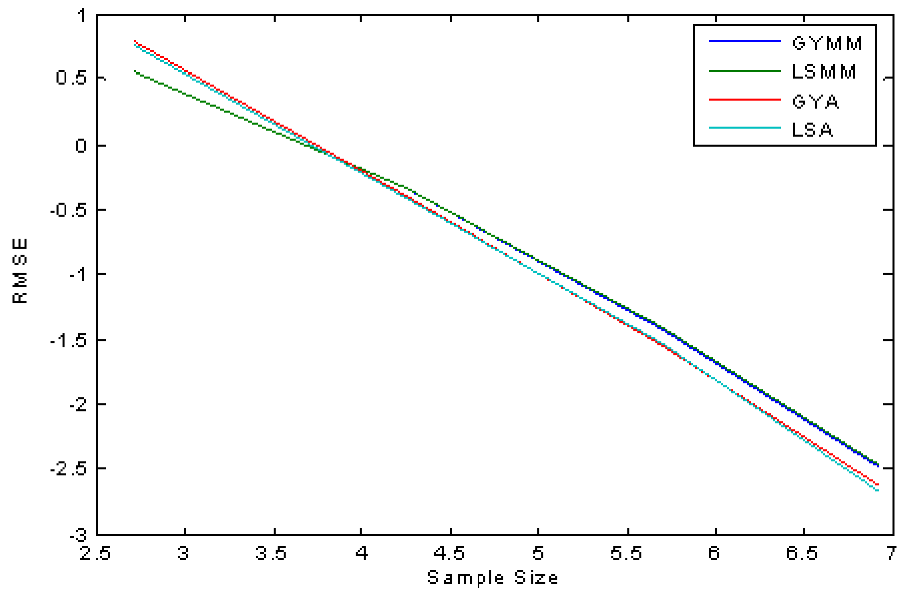

Figure 2: Root mean squares error versus sample size in pricing an American Put option with strike \$45 and initial stock price \$40.

VI Regression Methods and Control Variates

The method of control variates is one of the most popular variance reduction techniques and has many analogies with moment matching. Applications of this methodology in finance for pricing, (Rubinstein, et al, 1985), or model calibration (Glasserman and Yu, 2005) are very common. In this section we implement the Longstaff and Schwartz (2001) and Glasserman and Yu (2004b) methodologies using control variates.

Suppose that, given a stopping time $\tau \in \Gamma(t, T)$ and the state variable X_i, we want to estimate the price of an option that, as in (1), can be obtained by solving the following conditional expectation:

$$V_i(x) = \sup_{\tau \in \Gamma} E[\Theta(X_\tau)|X_i = x] \tag{13}$$

for the set of all possible stopping times τ.

Consider the functions $\phi_{ik}(x)$ and impose that

Definition 3 *For* $i = 1, ..., n-1, \phi_{ik}(x)$ *is in* $L^2(\phi(X_i))$*;* Π_i *denotes the orthogonal projection from* $L^2(\Omega)$ *onto the vector space generated by* $\{\phi_1(x), \phi_2(x), ..., \phi_K(x)\}$

Define the sample estimator of the option using M independent paths:

$$\Pi\widetilde{V}_i = \frac{1}{M}\sum_{j=1}^{\mathrm{M}} X_{.j} = V_i \tag{14}$$

Glasserman and Yu (2004a) show that under certain conditions the sample estimator of the option will converge almost surely to the option price. Define now the estimator of the option using control variates as follows:

$$\Pi\widehat{Z}_i = \Pi\widetilde{V}_i + \lambda_i[\Pi Y_i - E_i(Y)] \tag{15}$$

where λ is a previsible process in F with $E_F[(\lambda)]^2 < \infty$ and Y is a random variable for which we can compute the conditional expectation.

The sample estimator in (14) can be written as:

$$\Pi\widetilde{V}_i = \frac{1}{M}\sum_{j=1}^{M}(\widehat{Z}_{j.}) \tag{16}$$

$$\Pi\widehat{Z}_i = \Pi\widetilde{V}_i + \lambda_i[\Pi Y_i - E_i(Y)] \tag{17}$$

$$\lim_{i\to\infty} \lambda E[\Pi Y_\tau - E_i(Y))] = 0$$

Therefore the following result follows

$$E_i(\widehat{Z}) = V_i \tag{18}$$

From (15) it follows that $Var[Z_i(\lambda_i)]$, particularly we have $Var(\widehat{Z}_i) \leq Var(\widetilde{V}_i)$ if

$$\lambda_i^* = -\frac{Cov[\widehat{Z}_i Y_i]}{Var[Y_i]} \tag{19}$$

Therefore efficiency can be gained by minimizing λ_i in (19). One way of doing so is to use a simple Least Squares approach, that, we already use to compute estimates of the conditional expectation. The estimation of λ_i will introduce some bias. However, this will vanish as we increase the number of replications. As pointed out in Boyle, Broadie and Glasserman (1997), the estimator of λ_i does not need not be very precise to achieve a reduction of variance when only one control is used. It becomes important when multiple controls are introduced. In our empirical application we fix $\lambda_i = 1$.

VII Empirical Results

Table 3 shows an empirical application of control variate. To implement the estimator we sample the (discounted) price of a similar European option at each possible stopping time. This, by construction, should define a martingale at that time.

Results are rather encouraging and show that the prices estimated by control variates are rather precise, both with exponential and Laguerre basis. This result is in line with Broadie and Cao (2008). However in that study only few options with the same volatility are considered[6].

[6] Furthermore, we also consider different basis functions. Finally, it is worth mentioning that their numerical exercise assessing the Glasserman and Yu methodology is rather limited, while in this study is far more exhaustive.

		Exponential			Laguerre			Binomial
Basis Functions		2	3	4	2	3	4	
Methodolo	Vol./Mat.							
GY	0.3/0.333	5.691	5.705	5.702	0	0	0	5.706
SE		[0.00701]	[0.0089]	[0.0125]				
RMSE		[0.01536]	[0.0007]	[0.0044]				
LS	0.3/0.333	5.688	5.707	5.706	5.692	5.702	5.704	
SE		[0.0117]	[0.0089]	[0.0108]	[0.0076]	[0.0095]	[0.0088]	
RMSE		[0.01836]	[0.00046]	[0.00042]	[0.0144]	[0.0042]	[0.0017]	
GY	0.3/0.583	6.238	6.236	6.243	0	0	0	6.244
SE		[0.0098]	[0.0146]	[0.0096]				
RMSE		[0.0164]	[0.0077]	[0.0009]				
LS	0.3/0.583	6.229	6.241	6.242	6.226	6.239	6.246	
SE		[0.0125]	[0.0108]	[0.0139]	[0.00137]	[0.0161]	[0.0156]	
RMSE		[0.0154]	[0.0027]	[0.00175]	[0.0179]	[0.0005]	[0.0002]	
GY	0.4/0.333	6.491	6.507	6.513	0	0	0	6.51
SE		[0.0142]	[0.0151]	[0.01444]				
RMSE		[0.0201]	[0.00264]	[0.00359]				
LS	0.4/0.333	6.491	6.504	6.51	6.496	6.513	6.511	
SE		[0.0108]	[0.0131]	[0.0124]	[0.0128]	[0.0103]	[0.0105]	
RMSE		[0.01554]	[0.00593]	[0.00247]	[0.0144]	[0.0032]	[0.00049]	
GY	0.4/0.583	7.364	7.374	7.38	0	0	0	7.383
SE		[0.0158]	[0.0111]	[0.0123]				
RMSE		[0.0192]	[0.00873]	[0.00345]				
LS	0.4/0.583	7.363	7.381	7.382	7.365	7.377	7.383	
SE		[0.0101]	[0.01657]	[0.0145]	[0.01286]	[0.01336]	[0.0162]	
RMSE		[0.0202]	[0.00285]	[0.00133]	[0.01776]	[0.0058]	[0.000475]	

Table 3: Control Variates

Note that GY and LS are respectively the Longstaff-Schwartz and Glasserman-Yu methodologies. Zeros refer to cases when we were not able to implement the GY method. SE and RMSE are respectively standard errors and root mean square errors.

If we compare these results with the ones reported in Table 2, it is evident that the RMSEs are much smaller in this case. Therefore the error in the estimation of the conditional expectation (in Equation (6)) is also reduced. Standard errors are also in general smaller than the ones in Table 2. Consequently the small sample bias is also reduced. Note that when Laguerre basis are used the LS method shows the expected uniform convergence. Three basis are sufficient to achieve a low RMSE. It is not always the case that the GY (2004b) method produces the smallest standard errors. This result may not entirely support Theorem 1 in Glasserman and Yu (2004b)[7]. To

[7] The assumption of finite variance on the basis (see Assumption C1 in Glasserman and Yu, 2004b) may also be another reason. However a more detailed analysis is required here. This issue was also considered in Broadie and Cao (2008) reaching somehow a different conclusion. However, as already mentioned, the small sample investigation of the GY method is rather limited.

provide more empirical evidence supporting control variates, we measure the impact of control variate on the estimates of the option price in Table 3 by calculating the variance reduction factor (VR). The VR has been calculated as the ratio of the estimate of naive variance and the estimate of control variate variance. We have considered a sample of 30 options (including the options in Table 3) with volatilities 20-40% and time to expiry up to four months. The VR factor ranges between 0.45 and 5. If we choose the middle range, this result would imply that the variance of control variates estimators is 1/3 smaller than the variance of the Monte Carlo estimators. Obviously this is likely to have a substantial impact on the estimates of the continuation value and on the estimates of the option price. Therefore, practitioners wishing to use these methodologies to price American "plain vanilla" put options, should consider the Longstaff and Schwartz (2001) method implemented by control variates as showed in the previous section. Also, in line with previous studies we find that Laguerre basis are the ones suitable for these specific financial instruments.

VIII Valuing American Asian Options

We consider the previous methodologies when pricing more complex options such as American Asian options and options written on a maximum of $n-$assets. It is with this type of options that Monte Carlo methods become interesting.

As in Longstaff and Schwartz (2001) we consider pricing an American Asian option having also an initial lockout period. In order to use the option prices reported in Longstaff and Schwartz (2001) as benchmark, we consider an American call option that after an initial lock out period of three months can be exercised at any time up to maturity T. We assume $T = 2$ years. The average is the (continuous) arithmetic average of the underlying stock price calculated over the lock out period. We implement the LS (2001) and GY (2004b) methodologies by using control variates methods. There are no studies applying and implementing the GY (2004b) methodology in this context. The choice of the control in this case falls, obviously, on the price of an equivalent geometric option. Therefore, we use the methodology described above and the price of a geometric average option as a control. As in Longstaff and Schwartz (2001) the strike price is \$100, the risk free rate of interest 6% and volatility 20%. We use different scenarios for the stock price and assume 200 steps for both stock price and average. Results are reported in Table 4.

	Exponential	Laguerre	Exponential	Laguerre	Exponential	Laguerre
S/M	30,000		50,000		75,000	
80	0.921	0.9218	0.937	0.9451	0.9422	0.951
90	3.081	3.1061	3.211	3.312	3.321	3.312
100	7.492	7.522	7.679	7.845	7.8431	7.873
110	13.23	13.89	14.188	14.234	14.355	14.501
120	20.091	21.21	22.081	22.111	22.189	22.311

Table 4: American Asian options (LS, 2001).

S is the initial stock price and M the number of Monte Carlo simulations.

As in Longstaff and Schwartz (2001),we use the first eight Laguerre basis[8] and 30,000 to 75,000 replications. Furthermore we also consider exponential basis. Using finite difference methods LS (2001) report option prices equal to \$0.949 (\$80), \$3.267 (\$90), \$7.889 (\$100), \$14.538 (\$110) and \$22.423 (\$120)[9]. In general, our results support the ones in Tables 3 of Longstaff and Schwartz (2001). That is, the LS (2001) method produces a very accurate option price. If we calculate the early exercise values in this case and compare them with prices estimates reported in LS (2001) for the same options but using antithetic variates, we see that, with Laguerre basis and $m = 75,000$, differences in the early exercise values for the LS (2001) method ranges between 0.007 and 0.050, while in the present study the range is between 0.001 and 0.042. This is in line with what we pointed out at the beginning. The choice of variance reduction techniques is important when pricing option with American features since it reduces the probability of generating sub-optimal strategies.

[8]That is, first two Laguerre basis on the stock price and average plus cross products including an intercept.

[9]Number in the brackets are initial stock prices and the initial average value for the stock price is assumed to be 90.

In Table 5, we extend the Glasserman and Yu (2004b) method to price American Asian options. We use Hermite basis (ϕ_{KH}) to satisfy the Assumption 2, $f_K \phi_{KH}$ with $f_K = t^{K/2}$. The method seems to underestimate the option price.

S/M	30,000	50,000	75,000
80	0.925	0.936	0.94
90	3.188	3.311	3.166
100	7.521	7.544	7.563
110	13.83	14.223	14.321
120	20.11	21.645	22.022

Table 5: American Asian options (GY, 2004b).

S is the initial stock price while M is the number of Monte Carlo simulations.

However, in general, more work is necessary to implement this methodology since the choice of a martingale basis might be fundamental. On the other hand it seems that this fundamental problem has become more, to use Chris Rogers's words, "an art than a science." As pointed out above we shall address this important issue in a separate study. Therefore, in practical applications, as the one considered in this section, we suggest using the LS (2001) methodology and control variate in combination with Laguerre basis functions.

IX Valuing American Basket Options

Finally, we consider an additional high dimensional problem. We consider an American call option written on a maximum of five risky assets, paying a proportional dividend. We assume that each asset return is independent from the other. Once again, we use the same parameter specifications as in Longstaff and Schwartz (2001) and Broadie and Glaserman (1997) so that we can use prices reported in these papers as benchmark. In this application we simply use antithetic variates.

Broadie and Glasserman (1997) use stochastic mesh to price these type of options and report confidence intervals. We consider three different options with initial stock prices of 90,100, and 110 respectively. The assets pay a 10% proportional dividend, the strike price of the option is 100, the risk free rate of interest is 10%p.a and volatility 20%. Confidence intervals in Brodie and Glasserman (1997) are [16.602, 16.710] when the initial asset value is 90; [26.101, 26.211] with initial asset value of 100, and finally [36.719, 36.842] when the initial value is 110. The option prices in Longstaff and Schwartz (2001) are respectively, 16.657, 26.182, and 36.812 and they all fall within the Broadie and Glasserman 's confidence interval.

	Exponential	Hermite	Exponential	Hermite	Exponential	Hermite
S	M		M		M	
	30,000		50,000		75,000	
90	16.689	16.677	16.655	16.617	16.663	16.642
100	26.176	26.1744	26.171	26.103	26.081	26.121
110	36.769	36.764	36.783	36.748	36.8214	36.748

Table 6: American Basket Options (LS, 2001)

S is the stock price and M is the number of Monte Carlo replications

S/M	30,000	50,000	75,000
90	16.594	16.623	16.476
100	26.079	26.181	25.681
110	36.286	36.711	36.103

Table 7: American basket option (GY, 2004b).

S is the initial stock price while M is the number of Monte Carlo simulations.

All the estimated prices in Table 6 fall within Brodies and Glasserman's confidence intervals. We also extend the GY (2004b) method to price basket options (see Table 7)[10]. We use Hermite polynomials to satisfy Assumption 1 in GY (2004b). We note that option prices estimates fall within the Broadie and Glasserman 's confidence interval when 50,000 paths are considered. The martingale basis used in

[10] Again this is completely new in the literature.

this case seems to be appropriate. Therefore in practical applications we suggest using the LS (2001) methodology and Hermite basis. However, in this case, the Glasserman and Yu method, in combination with Hermite martingales basis, can also be used.

X Pricing American Derivatives under Stochastic Volatility

In this section we propose a novel approach, which extends the LS (2001) method to price options when the volatility is stochastic. This is a novel contribution in this area since the large majority of studies have only considered deterministic volatility. In this paper we consider the Heston (1999) stochastic volatility model to price American options. The Heston model is probably the most popular stochastic volatility model. The model is very popular because it can be extended to include, for example, jumps in a rather simple fashion. The variance process in this model follows a square root diffusion process

$$dS_t = \mu S_t dt + \sqrt{V_t} S_t dW_{1t} \tag{20}$$

$$dV_t = k(\theta - V_t)dt + \zeta\sqrt{V_t} dW_{2t} \tag{21}$$

$dW_{1t}dW_{2t} = \rho dt$ where $\sqrt{V_t}$ is the volatility, θ the long variance, k represents the speed of mean reversion, ζ is the volatility of variance, and ρ the correlation between stock returns and the volatility process. Finally W is a Winer process. Using the above model Heston (1999) derives a closed form solution for European options. For American options, as in the case of deterministic volatility, there is no closed form solution. However, stock prices can be simulated using Monte Carlo methods. An Euler discretization method consists in defining a set of times $t_0 < t_1 < ... < t_{n,}$ and using the Euler equation. This will introduce a discretization bias. However, this bias can be reduced by increasing the number of time steps. This will obviously reduce the efficiency of the method since it increases the computational effort. Therefore, the choice of the time steps to implement the Euler method is an important issue. In general, to achieve convergence of the simulated price to the true price, the number of times steps is set proportional to the square root of the number of simulations. But the constant of proportionality is rather difficult to determine in advance. Generally an Euler method applied to the variance process would use (22) instead of (21)

$$V_{t+1} = V(t_i) + k(\theta - V(t_i))[t_{i+1} - t_i] + \zeta\sqrt{V(t_i)}\sqrt{t_{i+1} - t_i}Z_{i+1} \tag{22}$$

where $Z_i \tilde{} N(0,1)$

In what follows we discuss a simple way of implementing this scheme. The main aspect of our simulation approach consists in transforming the model such that the discounted stock prices are martingales. There are different reasons for doing so. Firstly, this modification is necessary in order to preserve the martingale assumption, imposed upon many theoretical models. Also the constant expectation property of martingales might be important in order to control for the number of time steps in the simulation.

As in Heston the free arbitrage PDE that any option must satisfy in the presence of stochastic volatility is given by

$$\frac{\partial O}{\partial t} + \frac{1}{2}V^2S^2\frac{\partial^2 O}{\partial S^2} + \frac{1}{2}\zeta^2\frac{\partial O}{\partial V^2} + \rho VS\zeta\frac{\partial^2 O}{\partial S\partial V} - rO + rS\frac{\partial O}{\partial S} + (k[\theta - V]) - \Lambda(S,V,t)\zeta\sqrt{V}\frac{\partial O}{\partial V} = 0 \tag{23}$$

with the option price function given by $O = f(t, S_t, V_t)$ and $\Lambda(S,V,t)$ being the market price of risk.

By setting $\Lambda(S,V,t) = k\sqrt{V}$, or $\Lambda(S,V,t)\zeta\sqrt{V} = k\zeta V$, the above equation can be re-written as

$$\frac{\partial O}{\partial t} + \frac{1}{2}V^2S^2\frac{\partial^2 O}{\partial S^2} + \frac{1}{2}\zeta^2\frac{\partial O}{\partial V^2} + \rho VS\zeta\frac{\partial^2 O}{\partial S\partial V} - rO + rS\frac{\partial O}{\partial S} + (k[\theta - V]) - k\zeta V\frac{\partial O}{\partial V} = 0 \tag{24}$$

First note that Equation (24) can also be obtained within a risk neutral framework by means of a change of measure (i.e. moving from the original probability measure to the equivalent martingale measure). Also note that the term on $\frac{\partial O}{\partial V}$corresponds to the drift term in (22). This is therefore the risk neutralized drift term of the process (22). We propose a further transformation of the drift above. Set $\lambda = k\zeta$, then the market price of risk becomes $\Lambda(S, V, t) = \lambda V$, also set $k^* = k(1 + \zeta)$, and $\theta^* = \theta/k^*$, then the process in (22), after some algebra manipulation, can be re-written as

$$V(t_{i+1}) = V(t_i) + k^*(\theta^* - V(t_i))[t_{i+1} - t_i] + \zeta\sqrt{V(t_i)}\sqrt{t_{i+1} - t_i}Z_{i+1} \tag{25}$$

The drift term in (25) is now equivalent to the drift process in (24). In our empirical application, we have used this modified Euler equation to simulate the variance process. Finally, we use a reflection rule to avoid negative values for the process $\sqrt{V}$. Additionally to the variance process, to simulate the model we also need to simulate the stock price process. First note that the risk neutral drift on $\frac{\partial O}{\partial S}$ is also the risk neutralized drift term to be used in (20). With this transformation in mind and the one on the variance process, we have now that the discounted stock price process is a martingale. To implement the Euler method we apply the Ito Lemma to the portfolio process above and after some algebra manipulation we obtain the Euler equation for the stock process shown below

$$\ln(S)(t_{i+1}) = \ln(S(t_i) + (r - \tfrac{1}{2}V(t_i))[t_{i+1} - t_i] + \sqrt{V(t_i)}\sqrt{t_{i+1} - t_i}Z_{i+1}$$

There are several reason for using this approach. In fact, although under our drift transformation, discounted stock prices are martingales, we cannot be sure that these are positive martingales. To avoid a negative martingale process we consider the process above. Finally, the log transformation should further help us to reduce the discretization error.

XI Empirical Results

We first apply the methodology described above to price European options under stochastic volatility. In this case we can compute option prices in closed form using the Heston formula. To check that results are not driven by a specific choice of parameters, we consider a variety of different parameters and compute the absolute error with respect to the closed form price. We consider put options with strike equal to $10, initial stock price $9, the risk free rate of interest is 10%p.a., the long run mean variance is 0.16, the correlation coefficient 0.1, and the expiry three months. The basis functions considered are simple exponential basis (the first three) for the stock price, one exponential basis for the variance process and the cross product of the first basis for the stock price and the variance process. Prices have been computed using 50 times steps and 3000 simulations. Table 8 below reports the empirical result.

As expected our methodology can generate prices that are rather accurate. Given the modest number of time steps used in comparison with a standard Euler scheme approach, it is clear that there is a substantial gain in terms of computational time.

We now further extend the methodology to price American put options. We use the same specifications as above and control variates[11]. We use the approach described in Section VI, but the price of the European option is now sampled at maturity. Table 9 shows the empirical results.

	$V = 0.0625$		$V = 0.25$	
	$\zeta = 0.9$	$\zeta = 0.45$	$\zeta = 0.9$	$\zeta = 0.45$
	$k = 5$	$k = 2.5$	$k = 5$	$k = 2.5$
AE	0.0679	0.0284	0.0108	0.0263

Table 8: European Options Under Stochastic Volatility

Refer to Section X for notations. AE is the absolute error. Prices are averages of fifty trials.

[11] Its worth mentioning that the control variate method has not been considered in the literature (at least not on the best of our knowledge) in this area.

	$V = 0.0625$		$V = 0.25$	
	$\zeta = 0.9$	$\zeta = 0.45$	$\zeta = 0.9$	$\zeta = 0.45$
	$k = 5$	$k = 2.5$	$k = 5$	$k = 2.5$
S = 9				
MC	1.107	1.075	1.334	1.363
C(2008)	1.101	1.064	1.326	1.318
AE	0.006	0.01	0.008	0.045
SE	0.000102	0.00097	0.000187	0.000148
S = 10				
MC	0.521	0.478	0.793	0.837
C(2008)	0.518	0.4744	0.793	0.837
AE	0.0031	0.00357	0	0
SE	0.000396	0.000197	0.000934	0.000792

Table 9: American Options Under Stochastic Volatility

S is the initial stock price, MC and C (2008) refer to Monte Carlo and the method proposed in this paper.

We compare our approach with the Longstaff and Schwartz (2001) method-Monte Carlo (MC)- with 200,000 (antithetic variates) simulations and 500 time steps. Absolute errors (AE) gives an idea of the size of the error. This seems to be small for all the parameters we have chosen. Standard errors are also very small. We have also tested our methodology to price options with longer maturities. In fact some of the methodologies proposed in the literature fail exactly in this context. We only report one single case here. We consider an in-the-money put option with the same model specifications as the ones in the table above but one year to expiry and $V = 0.0625, \xi = 0.45, \kappa = 5$. The European option price is \$0.623. Using this price as a control, we compute the American option price absolute (0.0578) and standard errors (0.00065). Our methodology seems to be also applicable to price longer maturity options and it shows a reasonable degree of precision.

XII Conclusion

From an academic and even a practitioner‘s point of view, pricing American options still remains an interesting research area, particularly when Monte Carlo method is used. This is mainly due to the flexibility of this methodology to accommodate high dimensional features.

Recently, Longstaff and Schwartz (2001) and Glasserman and Yu (2004b) propose two option pricing estimators based on Monte Carlo simulations. This study contributes to the existing literature in many different ways. Firstly, it extends some recent estimators for American options pricing by implementing variance reduction techniques and shows that, in this way, one can reduce the probability of choosing sub-optimal exercise decisions and, consequently, reduce the option price bias. Rogers (2002) formulated the problem in Equation (3) as its dual and showed that one can use a martingale approach to reduce the probability of choosing sub-optimal policies when determining the early exercise value. However, that approach requires designing an optimal martingale and there is no clear cut rule yet on how to achieve it. In a companion paper we show how designing martingales bounded in the Hilbert space and pricing options under this measure. Of course our Monte Carlo analysis might be limited in the sense that it does not allow us to clearly define the best candidate (i.e. methodology). However, our empirical results may provide a guidance to traders. The methodology described in this study (see Table 3) produces standard errors and root mean squares errors that, in general, are of the same order of magnitude. This result implies that no error overcomes the other (i.e. small sample error and the error induced by using the Least Squares rule to determine the optimal stopping time). This should make our approach very appealing in empirical applications. This study also shows how extending the Glasserman and Yu (2004b) estimator to solve high dimensional problems. This is a novel contribution.

Finally, we propose a novel approach to price options under stochastic volatility. This represents a major contribution of this paper. The proposed methodology is flexible and efficient and it is compared with existing methods, providing in all cases precise option prices.

Overall, we found that option prices estimates using LS (2001) and GY (2004b) methodologies are accurate regardless the type of option considered (although for multidimensional problems the Longstaff and Schwartz, 2001 methodology and Laguerre basis seem to be the best combination). A large part of the sample bias can be eliminated with an acceptable number of replications (i.e. 100,000). With the Longstaff and Schwartz methodology the empirical results in Table 3 favour Laguerre polynomials and control variates. The choice of Laguerre basis is in line with Longstaff and Schwartz (2001). Therefore, in practical applications, we recommend using Laguerre polynomials. In general, a number of basis equal to three, 100,000 replication and control variate seem to be the right combination to achieve a substantial level of accuracy.

References

[1] Abadir, K., M., and P., Paruolo, 2008, "On Efficient Simulations in Dynamic Models", Tanaka Business School, Imperial College, Mimeo.

[2] Boyle, P., P., M., Broadie and P., Glasserman, 1997, "Monte Carlo Methods for Security Pricing", Journal of Economic, Dynamics and Control, 21, 1267-1321.

[3] Broadie, M. and P., Glasserman, 1997, "A stochastic Mesh Method for Pricing High Dimensional American Options, Working Papers, Columbia University.

[4] Broadie, M., and O., Kaya, 2004, "Exact Simulation of Stochastic Volatility and other Affine Jump Diffusion Processes", Columbia University, Graduate School of Business.

[5] Broadie, M., and M., Cao, 2008, "Improved Lower and Upper Bound algorithms for Pricing American Options by Simulation", Quantitative Finance, 8, Vol 8.

[6] Clement, E., D., Lamberton and P., Protter, 2002, "An Analysis of a Least Squares Regression Method for American Option Pricing", Finance and Stochastic, 6, 449-471.

[7] Cox, J., C., S.A.Ross and M.,Rubinstein, 1979, “Option Pricing:A Simplified Approach”, Journal of Financial economics, 7, 229-263.

[8] Egloff, D., 2003, “Monte Carlo Algorithms for Optimal Stopping and Statistical Learning”, Working Paper, Zurich Kantonalbank, Zurich Switzerland.

[9] Glasserman, P., 2004, "Monte Carlo Methods in Financial Engineering", Springer.

[10] Heston, S., 1999, “A closed-form solution for options with stochastic volatility with applications to bond and currency options, Review of Financial Studies 6, 327-343.

[11] Glasserman, P., and B., Yu, 2004a, “Number of Paths Versus Number of Basis Functions in American Option Pricing”, The Annals of Applied Probability, Vol.14.

[12] Glasserman, P., and B., Yu, 2004b, “Simulation for American Options:Regression Now or Regression Later?”, in H., Niederreiter editor, Monte Carlo and Quasi Monte Carlo Methods, 2002.

[13] Glasserman, P., and B., Yu, 2005, " Large Sample Properties of Weighted Monte Carlo Estimators", Operation Research, 53, 298-312.

[14] Longstaff, F., A., and E., S., Schwartz, 2001,”Valuing American Options by Simulation: A SimpleLeast-Squares Approach”, The Review of Financial Studies, Vol. 14, No.1.

[15] Rasmussen, N., 2005, "Control Variate for Monte Carlo Valuation of American Options", Journal of Computational Finance, 9, Vol1.

[16] Rogers, L., C., 2002, “Monte Carlo Valuation of American Options”, Mathematical Finance, 12, 271-286.

[17] Rubinstein, R., Y., and R., Marcus, 1985, “Efficiency of Multivariate Control Variates in Monte Carlo Simulations”, Operations Research, 33.

[18] Stentoft, L., 2004, “Assessing the Least Squares Monte Carlo Approach to American Option Valuation”, Review of Derivatives research, 7(3).