Journal of
Risk and Financial Management

MDPI

Article

Efficient Variance Reduction for American Call Options Using Symmetry Arguments

François-Michel Boire [1,*], R. Mark Reesor [2] and Lars Stentoft [1,3]

[1] Department of Statistical and Actuarial Sciences, University of Western Ontario, London, ON N6A 5B7, Canada; lars.stentoft@uwo.ca
[2] Department of Mathematics, Wilfrid Laurier University, Waterloo, ON N2L 3C7, Canada; mreesor@wlu.ca
[3] Department of Economics, University of Western Ontario, London, ON N6A 5C2, Canada
* Correspondence: fboire2@uwo.ca

Abstract: Recently it was shown that the estimated American call prices obtained with regression and simulation based methods can be significantly improved on by using put-call symmetry. This paper extends these results and demonstrates that it is also possible to significantly reduce the variance of the estimated call price by applying variance reduction techniques to corresponding symmetric put options. First, by comparing performance for pairs of call and (symmetric) put options for which the solution coincides, our results show that efficiency gains from variance reduction methods are different for calls and symmetric puts. Second, control variates should always be used and is the most efficient method. Furthermore, since control variates is more effective for puts than calls, and since symmetric pricing already offers some variance reduction, we demonstrate that drastic reductions in the standard deviation of the estimated call price is obtained by combining all three variance reduction techniques in a symmetric pricing approach. This reduces the standard deviation by a factor of over 20 for long maturity call options on highly volatile assets. Finally, we show that our findings are not particular to using in-sample pricing but also hold when using an out-of-sample pricing approach.

Keywords: antithetic sampling; control variates; importance sampling; Monte Carlo simulation; put-call symmetry

JEL Classification: C15; G12; G13

check for updates

Citation: Boire, François-Michel, R. Mark Reesor, and Lars Stentoft. 2021. Efficient Variance Reduction for American Call Options Using Symmetry Arguments. *Journal of Risk and Financial Management* 14: 504. https://doi.org/10.3390/jrfm14110504

Academic Editor: S. Ejaz Ahmed

Received: 16 August 2021
Accepted: 9 October 2021
Published: 20 October 2021

1. Introduction

Key to valuing American options with a dynamic programming approach is estimating a continuation value that determines the optimal exercise strategy. For this task, simulation and regression-based methods are nowadays often preferred to other deterministic algorithms, like finite differences and multinomial trees, because they are easy to implement and because of their flexibility. The Least-Squares Monte Carlo (LSM) method of Longstaff and Schwartz (2001) is particularly popular, and several arguments have been made in favor of this methodology for valuing American options (Stentoft 2014). However, like other Monte Carlo pricing methods the LSM method is numerically costly and reducing its variance is therefore important.

This paper examines the relationship between the efficiency of variance reduction techniques and option features like moneyness, maturity, and asset volatility when pricing American-style options with the LSM method. Whereas most of the American option pricing literature has focused on either put or call options individually, we employ the symmetry relation of McDonald and Schroder (1998) such that we can readily compare results for pairs of call and put options whose solutions coincide. Three classical variance reduction techniques are studied in the context of LSM pricing: antithetic sampling, control variates, and importance sampling. We also consider implementations where two or more variance reduction techniques are combined. We restrict our attention to these three techniques because of their popularity, and because they do not require simulating additional paths. These estimators retain the flexibility of a crude LSM implementation and can easily be combined with moment-matching simulation (MMS) of Barraquand and Martineau (1995), or empirical martingale simulation (Duan and Simonato 1998). Other notable variance reduction techniques include low-discrepancy sequences (Lemieux and La 2005), latin hypercube (Glasserman 2003), and stratified sampling (Glasserman et al. 1999). If several estimators are readily available, optimal linear combinations of these estimators can also be explored.

Our results first show that efficiency gains from variance reduction may be quite different for calls and symmetric puts. However, control variates is by far the most efficient of the three methods. We observe that efficiencies, defined as the ratio of the standard deviations of the crude LSM and the LSM with variance reduction, increase (or never worsen) with time to maturity and asset volatility for symmetric put options. Conversely, efficiencies decrease (or never improve) with time to maturity and asset volatility for call options. Next, since control variates is always more effective for puts than calls, and since symmetric pricing already offers some variance reduction, we demonstrate that a drastic reduction of the standard deviation of the call option price is obtained by combining variance reduction techniques with a symmetric pricing approach. This is particularly so for long maturity call options on volatile assets for which the standard deviation can be reduced by a factor of over 20 when combining all three variance reduction techniques. Finally, we show that these results continue to hold when using an out-of-sample pricing approach.

There are other contributions to the literature on variance reduction techniques for the LSM method. For example, antithetic sampling was used by Longstaff and Schwartz (2001). Control variates was used with the LSM method as early as in Tian and Burrage (2002), and an optimal implementation was suggested in Rasmussen (2005). Importance sampling techniques were discussed in the context of the LSM method in Moreni (2003), and the selection of an optimal importance density was analyzed further in Bolia et al. (2004), Juneja and Kalra (2009), and Morales (2006).[1] However, to our knowledge, the present paper is the first to thoroughly study the simultaneous combination of variance reduction techniques for the Monte Carlo valuation of American-style options with the LSM method and to compare the results for pairs of call and symmetric put options whose solutions coincide such that relative efficiencies can be readily compared.

Our findings have important implications for the potentially complicated problem of efficiently pricing American-style call options. For the 40 American call options considered herein, the implementation of standard variance reduction techniques together with the symmetric pricing approach results in a drastic reduction of the variance over the crude LSM estimator. In particular, combining the symmetric pricing approach with combinations of variance reduction techniques with the control variates method largely facilitates the valuation of long term American call options written on volatile assets.

Our suggested approach is therefore of particular relevance for the valuation of real options, which often take the form of costly investment opportunities for projects whose life may span several decades, and underlying assets are difficult to model at long horizons. Chapter 22 of Brealey et al. (2018), for example, presents several examples of such real call options. In all these cases, the adequate valuation of real options is of paramount importance to give full financial flexibility to a firm's operations. Indeed, the value of the early-exercise feature of such investment opportunities can be seen as the "American premium" of the option, which is most prominent when the option is deep out-of-the-money (OTM), has long maturity, and when the underlying asset processes are volatile or difficult to predict reliably over the life of the option. Moreover, it is worth noting that when the symmetry property holds, a bijective relationship between the exercise boundaries of call and symmetric put options (Detemple 2001) informs practitioners not only about option prices, but also about optimal stochastic control, which is key to the management of real options.

The paper is organized as follows: Section 2 outlines the American option pricing problem, provides details on the implementation of the LSM method, and presents the put-call symmetry relation. Section 3 presents the results of our numerical experiments, compares variance reduction for comparable call and put options and discusses the efficiency gains made possible by the put-call symmetry. Finally, Section 4 concludes. Appendix A describes the variance reduction techniques and discusses some numerical issues that may arise when implementing them. Some additional figures are presented in Appendix B.

2. Pricing Derivatives with Early-Exercise Features

In this section we first state the valuation problem associated with pricing American options. Next, we demonstrate how the price can be approximated using the Least Squares Monte Carlo method of Longstaff and Schwartz (2001). Finally, we review the put-call symmetry property and present some initial numerical results.

2.1. The Valuation Problem

Consider an American option written on an underlying asset $S(t) : t \in [0, T]$ defined on a continuous filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t=0}^T$, where $\mathbb{P}$ is a risk-neutral or pricing probability measure. Consistent with the notation of Glasserman (2003), the objective of this problem is to maximize the discounted payoff $U(t) : t \in [0, T]$ with respect to an $\mathbb{F}$-adapted class of stopping times $\mathcal{T} : \mathcal{T} \subseteq [0, T]$. With underlying asset value $S(t)$, necessarily adapted to $\mathbb{F}$ and constant continuously-compounded interest rate r, we pose the optimal stopping time problem as

$$\begin{aligned} U(0) &= \sup_{\tau \in \mathcal{T}} \mathbb{E}[U(\tau) \mid \mathcal{F}_0] \\ &= \sup_{\tau \in \mathcal{T}} \mathbb{E}[e^{-r\tau} h(S(\tau)) \mid \mathcal{F}_0] \\ &= \mathbb{E}[U(\tau^*) \mid \mathcal{F}_0], \end{aligned} \tag{1}$$

where $h(\cdot) \geq 0$ is an $\mathbb{F}$-adapted payoff function, and τ^* is the optimal $\mathbb{F}$-adapted stopping time. Note that for notational simplicity, we have expressed the payoff function as depending only on $S(\tau)$. For instance, the value of an American put option with strike K and payoff $h(S(t)) = (K - S(t))^+$ is given by

$$P(0) = \sup_{\tau \in \mathcal{T}} \mathbb{E}\left[e^{-r\tau}(K - S(\tau))^+ \mid \mathcal{F}_0\right]. \tag{2}$$

In general, the payoff need only be adapted to the filtration.

For a put option, the problem of maximizing the expected discounted payoff with respect to a stopping time τ can be seen as the dual problem to the minimization (primal) problem given by

$$\tau^* = \inf\{\tau \in \mathcal{T} : S(\tau) \leq b^*(\tau)\}, \tag{3}$$

where $b^*(t)$ represents the optimal exercise boundary at time t. That is, whenever the underlying asset price goes below the threshold, immediate exercise is optimal. Otherwise, the option should be held. Similarly, a call option should be exercised when the underlying asset price is above its optimal exercise boundary. Thus, each exercise strategy corresponds to a stopping time and determines the American option price via Equation (1).

Let us consider a discrete-time formulation of the problem. For an option with a maturity of T years, we consider an evenly spaced partition with J time steps of length $\Delta t = T/J$. The discretized price process of the underlying asset $\{S_j : j = 0, \ldots, J\}$ is defined on a complete risk-neutral probability space $(\Omega, \mathcal{G}, \mathbb{P})$ and adapted to the discrete filtration $\mathbb{G} = \{\mathcal{G}_j\}_{j=0}^J$. For simplicity, we use time j to refer to time $\tau_j = j\Delta t$. Option exercise is allowed only at the points defining the time discretization, hence the set of admissible exercise opportunities is $\{\tau_j = j\Delta t : j = 0, \ldots, J\}$, where $0 = \tau_0 < \tau_1 < \cdots < \tau_J = T$. Let $V_j(x)$ denote the time-j value of an unexercised option with underlying asset value $S_j = x$. The option price can then be written as $V_0(S_0)$, which solves the dynamic programming recursion

$$\begin{cases} V_J(S_J) = h(S_J) \\ V_j(S_j) = \max\Big(h(S_j), \mathbb{E}\Big[e^{-r\Delta t} V_{j+1}(S_{j+1}) \mid \mathcal{G}_j\Big]\Big), j = J-1, \dots, 0. \end{cases} \tag{4}$$

Options with a discrete early exercise feature of this sort are termed Bermudan options. The continuously exercisable American option price is approximated by letting N tend to infinity (Glasserman 2003). In this paper we work with the time-discretized problem, and describe the options as American rather than Bermudan.

2.2. Least-Squares Monte Carlo

Most American option valuation methods rely on the dynamic programming representation in Equation (4) to obtain price estimates. The problem remains to estimate the time-j continuation value of the option, i.e., the quantity $\mathbb{E}\big[e^{-r\Delta t} V_{j+1}(S_{j+1}) \mid \mathcal{G}_j\big], j = 0, \dots, (J-1)$. Since the conditional expectation of a square-integrable function relative to a sigma algebra $\mathcal{G}$ can be represented as a countable linear combination of $\mathcal{G}$-measurable basis functions $\{\psi_l(\cdot)\}_{l=0}^{\infty}$ (see Royden 1988), the time-j continuation value can be written as

$$\mathbb{E}\Big[e^{-r\Delta t} V_{j+1}(S_{j+1}) \mid \mathcal{G}_j\Big] = \sum_{l=0}^{\infty} \psi_l(S_j)\gamma_{j,l}, \tag{5}$$

with associated real coefficients $\{\gamma_{j,l}\}_{l=0}^{\infty}$.

However, in practise this quantity cannot be computed and a finite number of basis functions is used. Given a set of N simulated paths $\{S_{n,j} : n = 1, \dots, N; j = 0, \dots, J\}$, the LSM method resorts to a parametric approximation of the continuation value with a regression model of the form

$$e^{-r\Delta t} V_{j+1}(S_{n,j+1}) = \sum_{l=0}^{L} \psi_l(S_{n,j})\beta_{j,l} + \varepsilon_{n,j+1}. \tag{6}$$

The $(L+1) \times 1$ vector of coefficient estimates $\hat{\beta}_j = \{\hat{\beta}_{j,0}, \dots, \hat{\beta}_{j,L}\}'$ are obtained by regressing disounted cashflows $e^{-r\Delta t} V_{j+1}(S_{n,j+1})$ against the cross-section of basis functions that relate to time-j in-the-money (ITM) asset paths, denoted by the $1 \times (L+1)$ vector $\psi(S_{i,j}) = \{\psi_0(S_{n,j}), \dots, \psi_L(S_{n,j})\}$, with ψ_0 as a constant. The time-j path-n fitted continuation value then takes the form

$$\hat{C}_{n,j} = \psi(S_{n,j})\hat{\beta}_j, \tag{7}$$

and determines an exercise strategy in which the option is exercised if the payoff is positive and greater than the fitted continuation value. The continuation value approximation of order L provides an estimate of the exercise strategy $b^L(t) : t \in \mathcal{T}$ with a corresponding stopping time τ^L determined by the choice of $L < \infty$. To simplify notation, we hereafter refer to $\tau(n)$ as the path-n LSM stopping time estimate.

The LSM estimator of Longstaff and Schwartz (2001) computes discounted cashflows along each path with the regression approach outlined above. To reduce the notation, we write $\hat{V}_j(S_{n,j})$ as $\hat{V}_{n,j}$ to obtain the *approximate* dynamic program

$$\begin{cases} \hat{V}_{n,J} = h(S_{n,J}) \\ \hat{V}_{n,j} = \begin{cases} h(S_{n,j}) & \text{if } (h(S_{n,j}) \geq \hat{C}_{n,j}) \cap (h(S_{n,j}) > 0) \\ e^{-r\Delta t}\hat{V}_{n,j+1} & \text{if } (h(S_{n,j}) < \hat{C}_{n,j}) \cup (h(S_{n,j}) = 0) \end{cases} \quad j = J-1, \dots, 0. \end{cases} \tag{8}$$

In this dynamic programming representation, the terminal option value is again set to the option payoff at maturity, specifying a continuation value for the next iteration of the backwards-in-time recursion. At the penultimate time step, the criterion for the path-n

exercise decision is that the payoff $h(S_{n,J-1})$ is positive and greater than the continuation value $\hat{C}_{n,J-1}$. If the option is exercised, the option value is the immediate payoff. Otherwise, the pathwise option cashflow is discounted back one time step. The algorithm then moves backwards in time, computing option values along all paths at each time, and updating the exercise decisions. If the exercise value never exceeds the continuation value along a simulated path, the option is unexercised, its payoff null, and hence the option value for that path is zero. If the payoff is positive and exceeds the continuation value at least once, the $\mathbb{G}$-adapted stopping time determined by the LSM is the first exercise time.

2.3. Put-Call Symmetry

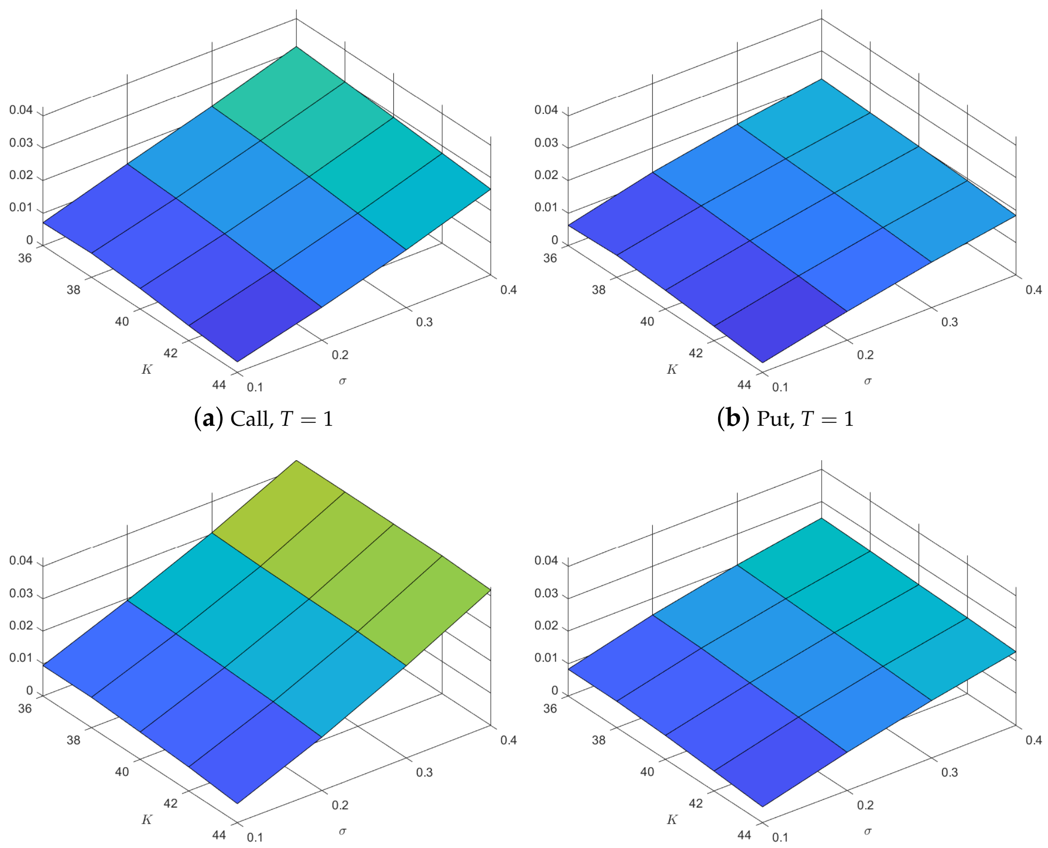
Throughout this paper, results are presented for pairs of American call and put options linked by a put-call symmetry result. Consider an American option with strike price K and maturity T, written on an underlying asset governed by a Geometric Brownian Motion and with initial price S_0, volatility of returns σ, and continuous interest rate r and dividend yield q. Denoting put and call option prices by $P(S_0, K, r, q, \sigma, T)$ and $C(S_0, K, r, q, \sigma, T)$, respectively, put-call symmetry states that the following equality holds

$$C(S_0, K, r, q, \sigma, T) = P(K, S_0, q, r, \sigma, T). \tag{9}$$

This result was first presented by McDonald and Schroder (1998). By systematically comparing symmetric put and call prices over a wide range of time to maturity, moneyness, and asset volatility, we can readily interpret and compare the efficiencies of variance reduction tools for pairs of problems whose solutions coincide.

Recent literature demonstrates that the symmetry relation is useful because the valuation of American call options poses a number of numerical problems which are otherwise not encountered for put options. In particular, because the payoff of a call option is unbounded, the presence of highly volatile paths will result in a higher frequency of deep moneyness in the cross-section of asset paths. This can impede the approximation of a decision rule enough to exacerbate variance and create biased price estimates. Symmetric pricing stands as a costless and effective way to value call options and the technique proves particularly effective in cases of long maturity options written on volatile assets.

As an illustration, consider call and (symmetric) put options with 50 exercise opportunities per annum, an initial underlying price $S_0 = 40$, risk-free rate $r = 0.06$, dividend yield $q = 0.06$, and maturities of 1 and 2 years. Figure 1 shows the standard deviation calculated from $M = 1000$ replications of an LSM configuration with $N = 100{,}000$ paths and a cubic approximation $(L = 3)$ of the decision rule. The figure shows that the standard deviation is indeed reduced when using symmetric pricing and this is particularly so for longer maturity options. For shorter maturities and less volatile assets, the symmetric pricing approach has subdued efficiencies in terms of standard deviations, but never produces significantly worse price estimates.



(**a**) Call, $T = 1$ (**b**) Put, $T = 1$

(**c**) Call, $T = 2$ (**d**) Put, $T = 2$

Figure 1. Standard deviation for symmetric options. Standard deviation results are calculated from $M = 1000$ replications of a standard LSM configuration with $N = 100{,}000$ paths and a cubic approximation ($L = 3$) of the decision rule. The left column illustrates price estimates for call options, and the right column for symmetric put options. The top and bottom rows present results for one- and two-year options, respectively.

3. Empirical Results

The Least-Squares Monte Carlo approach for pricing American derivatives has been shown to converge to the true price when the number of simulated paths, N, and the order of the polynomial expansion for the basis functions, L, tend to infinity (Stentoft 2004). In any actual implementation though, a finite number for both is used. In our numerical implementation of the LSM we use $L = 3$ basis functions and $N = 100{,}000$ sample paths in total to estimate the decision boundary.[2] We price options with $J = 50T$ (i.e., 50 exercise opportunities per annum), $r = q = 0.06$, and $S_0 = 40$. The numerical experiment carried out here consists in pricing 40 call options with varying levels of volatility, $\sigma = \{0.1, 0.2, 0.3, 0.4\}$, different strike prices, $K = \{36, 38, 40, 42, 44\}$, and maturities $T = \{1, 2\}$, as well as their 40 symmetric put counterparts. Results are based on $M = 1000$ independent replications.

We first compare the variance reduction that can be achieved for call and (symmetric) put options, respectively, and demonstrate that (1) efficiency gains are different for calls and symmetric puts, and (2) control variates should always be used. Second, we note that when using control variates, either alone or in combination with other techniques, efficiencies are always larger for put options than for the corresponding call options and we demonstrate that the joint or total effect of using (combinations of) variance reduction techniques together with symmetric pricing for call options can lead to price estimates with substantially lower variance. We demonstrate that these results are qualitatively identical when using an out-of-sample LSM approach in which the approximated decision rule is applied to a second independent sample of simulated paths. Finally, we provide some intuition for why variance reduction techniques work better than others when it comes to pricing American call and put options.

3.1. Variance Reduction for Call and Put Options

We start by examining the variance reduction that can be achieved by implementing three stand-alone techniques and combinations thereof for call and (symmetric) put options. Using the symmetric put options allows a direct comparison of the variance reduction for call and put options with similar characteristics and prices. The methods considered are antithetic sampling, denoted with an "A", control variates, denoted with a "C", and importance sampling, denoted with an "I". To compare performance, we consider the standard deviation efficiencies defined as

$$\mathrm{Eff}_{VR} = \sqrt{\frac{\mathrm{Var}\left[\bar{V}^{(N)}\right]}{\mathrm{Var}\left[\bar{V}_{VR}^{(N)}\right]}} - 1, \tag{10}$$

where $\bar{V}^{(N)}$ is the crude LSM estimator of an American option using a sample of N paths. Similarly $\bar{V}_{VR}^{(N)}$ is the estimator supplemented with a (combination of) variance reduction technique(s).[3] For example, $\bar{V}_{ACI}^{(N)}$ is the price estimator that combines all three variance reduction techniques, and Eff_{ACI} denotes its efficiency. The variances are estimated using the $M = 1000$ replications. Efficiencies greater than zero indicate improvement over crude Monte Carlo.

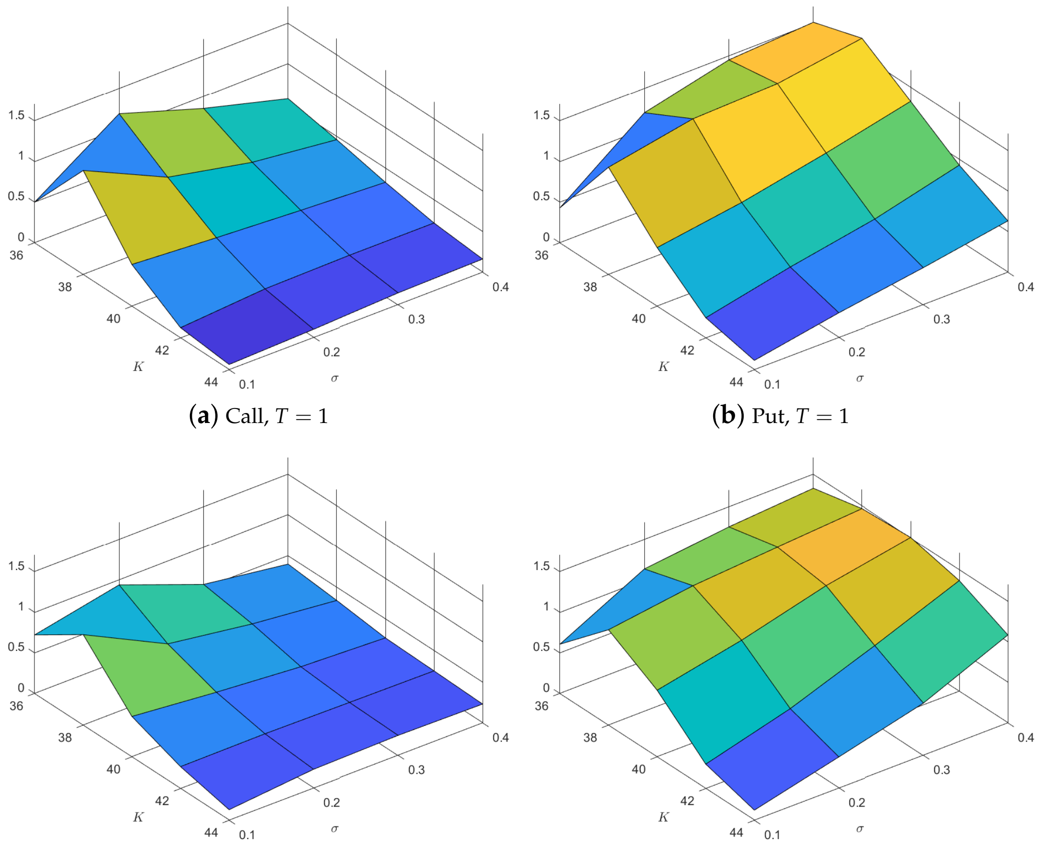
Table 1 reports the results for each of the variance reduction methods with efficiencies for call options in columns 4–10, and symmetric put options in columns 11–17. Option characteristics are shown in the first three columns of Table 1. The last row in the table reports the average efficiency for a given variance reduction technique across the call and put options, respectively. The table first demonstrates that efficiency gains from variance reduction techniques are different for call and put options across option characteristics. For example, and we highlight this in Figure 2, when using only antithetic sampling the efficiency increases with moneyness and is insensitive to variations in volatility for call options, whereas the efficiency is highest for at-the-money (ATM) options and increases with the volatility of the underlying asset for put options. Across the methods we observe that efficiencies increase (or never worsen) with time to maturity and asset volatility for symmetric put options whereas efficiencies decrease (or never improve) with time to maturity and asset volatility for call options. For all the individual options as well as on average, some efficiency gain is achieved as all efficiencies are positive.

Secondly, the table shows that the best performing variance reduction method, i.e., the method with the highest efficiency, always involves the use of control variates. As a result the average efficiencies obtained with control variates are much larger than those obtained with antithetic or importance sampling, and in the case of put options this is so by an order of magnitude. The ACI-LSM approach, which combines all three stand-alone variance techniques, is the most efficient approach for most of the options and for 36 out of the 40 put options. ACI-LSM is also the most efficient method for 17 out of the 20 long maturity call options (i.e., options with a maturity of 2 years) and when it is not the best perfoming method it has efficiencies that are very close to the best estimators. In fact, the methods that involve control variates all perform very similarly for call options with very small relative differences and close to optimal variance reductions can be achieved with the computationally simple stand-alone control variate method. For put options however, adding antithetic and importance sampling does further reduce the variance by $(14.1 - 12.4)/12.4 = 13.7\%$ demonstrating that combining all three variance reduction techniques does improve on the efficiency for these options. Figures A3 and A4 in Appendix B provide a visual representation of this.

Table 1. Efficiency results for various (combinations of) variance reduction techniques.

			Call							**Symmetric Put**						
T	K	σ	**A**	**C**	**I**	**AC**	**AI**	**CI**	**ACI**	**A**	**C**	**I**	**AC**	**AI**	**CI**	**ACI**
1	36	0.1	0.5	8.5	1.1	**9.0**	1.3	7.9	8.9	0.4	10.0	0.9	10.5	1.1	9.5	**10.6**
1	36	0.2	1.2	**7.6**	1.4	7.2	1.6	7.4	7.4	1.2	11.2	1.1	11.6	1.3	11.1	**12.8**
1	36	0.3	0.9	**6.8**	1.6	6.6	1.9	6.8	6.7	1.5	11.4	1.1	12.8	1.3	11.4	**14.4**
1	36	0.4	0.6	6.1	1.8	5.7	2.2	**6.2**	5.8	1.5	12.5	1.1	13.4	1.2	12.6	**15.1**
1	38	0.1	1.3	8.7	1.3	8.7	1.6	8.5	**9.1**	1.3	10.5	1.2	11.1	1.4	10.3	**12.1**
1	38	0.2	0.8	**8.1**	1.5	7.9	1.8	8.1	7.9	1.5	11.1	1.3	12.3	1.4	11.1	**13.2**
1	38	0.3	0.6	7.2	1.7	7.2	2.1	**7.3**	7.3	1.5	11.9	1.2	13.2	1.3	12.0	**14.1**
1	38	0.4	0.5	6.0	1.8	5.9	2.2	**6.1**	6.0	1.7	12.2	1.2	13.6	1.3	12.4	**14.8**
1	40	0.1	0.5	9.2	1.6	**9.4**	1.9	9.2	9.4	0.7	10.8	1.5	11.8	1.7	10.8	**11.8**
1	40	0.2	0.4	8.1	1.7	8.5	2.1	8.1	**8.6**	0.9	11.8	1.4	13.0	1.6	11.9	**13.1**
1	40	0.3	0.4	7.2	1.8	7.3	2.3	7.3	**7.3**	1.1	12.2	1.3	13.4	1.5	12.4	**13.7**
1	40	0.4	0.3	6.0	1.9	5.6	2.3	**6.0**	5.6	1.3	12.7	1.3	14.0	1.4	12.9	**14.5**
1	42	0.1	0.1	8.8	2.0	8.8	2.3	**8.9**	8.8	0.2	11.8	2.1	11.3	2.2	**11.9**	11.3
1	42	0.2	0.2	7.8	1.8	7.9	2.3	7.9	**8.0**	0.5	12.5	1.7	12.6	1.8	12.5	**12.6**
1	42	0.3	0.2	7.0	1.9	6.8	2.4	**7.1**	6.8	0.7	12.8	1.5	13.2	1.7	12.9	**13.3**
1	42	0.4	0.2	5.9	1.9	5.5	2.4	**6.0**	5.5	0.9	13.3	1.4	14.0	1.5	13.6	**14.2**
1	44	0.1	0.1	10.3	2.9	10.0	2.8	**10.4**	10.2	0.1	11.4	2.8	12.2	2.6	11.6	**12.3**
1	44	0.2	0.1	8.2	2.2	7.8	2.4	**8.3**	7.9	0.3	**13.0**	1.9	12.3	2.1	13.0	12.3
1	44	0.3	0.1	6.9	2.0	6.7	2.5	**7.0**	6.7	0.5	13.8	1.6	13.4	1.8	**13.8**	13.5
1	44	0.4	0.2	5.9	2.0	5.6	2.5	**6.0**	5.7	0.6	13.7	1.5	13.9	1.6	13.9	**14.1**
2	36	0.1	0.7	9.6	1.8	9.6	1.8	9.8	**11.3**	0.6	10.0	1.4	9.8	1.4	10.2	**11.6**
2	36	0.2	0.9	8.1	2.0	7.7	1.9	8.5	**8.7**	1.1	11.3	1.4	11.2	1.4	12.5	**16.1**
2	36	0.3	0.5	6.4	2.2	6.4	2.2	6.5	**6.8**	1.3	11.9	1.3	11.6	1.3	13.5	**16.5**
2	36	0.4	0.4	4.0	2.1	4.0	2.1	4.1	**4.1**	1.3	12.6	1.2	11.8	1.2	14.3	**16.3**
2	38	0.1	1.1	9.4	1.9	8.8	1.8	9.9	**10.6**	1.2	11.4	1.5	10.6	1.5	12.4	**14.8**
2	38	0.2	0.6	7.9	2.1	8.1	2.1	8.3	**8.7**	1.3	12.9	1.5	11.5	1.5	14.1	**14.6**
2	38	0.3	0.4	6.2	2.2	6.4	2.3	6.3	**6.6**	1.4	13.3	1.3	11.8	1.3	14.5	**15.2**
2	38	0.4	0.3	3.9	2.1	4.1	2.2	4.0	**4.2**	1.5	13.7	1.2	11.9	1.2	15.0	**15.6**
2	40	0.1	0.5	9.3	2.1	9.9	2.2	9.7	**10.4**	0.8	11.9	1.7	12.7	1.8	12.3	**13.8**
2	40	0.2	0.4	7.9	2.3	8.7	2.4	8.1	**9.0**	1.0	12.9	1.5	12.8	1.5	13.5	**14.4**
2	40	0.3	0.3	5.9	2.3	6.4	2.5	6.1	**6.6**	1.3	13.4	1.3	12.4	1.4	14.3	**14.6**
2	40	0.4	0.3	4.0	2.2	4.1	2.3	4.1	**4.2**	1.5	13.9	1.2	12.8	1.2	15.1	**15.9**
2	42	0.1	0.3	10.2	2.4	10.3	2.8	10.4	**10.4**	0.3	12.4	1.9	13.2	2.1	13.0	**13.6**
2	42	0.2	0.3	7.9	2.4	8.8	2.8	8.1	**9.1**	0.6	13.4	1.6	14.0	1.7	14.3	**15.3**
2	42	0.3	0.3	6.0	2.5	6.7	2.8	6.2	**6.8**	1.0	13.4	1.4	14.2	1.4	14.2	**15.9**
2	42	0.4	0.3	4.2	2.3	4.2	2.5	4.2	**4.3**	1.4	14.0	1.3	13.4	1.3	14.8	**15.7**
2	44	0.1	0.1	9.8	2.6	9.4	3.2	**10.1**	9.6	0.1	11.9	2.2	11.8	2.5	**12.4**	12.1
2	44	0.2	0.2	8.3	2.6	8.5	3.0	8.4	**8.6**	0.4	13.5	1.7	14.6	1.8	14.5	**15.3**
2	44	0.3	0.2	6.3	2.6	6.1	2.9	**6.4**	6.2	0.6	13.7	1.5	14.3	1.5	14.8	**15.6**
2	44	0.4	0.2	4.3	2.4	4.1	2.6	**4.3**	4.2	1.1	14.0	1.3	14.6	1.3	14.9	**16.8**
Average			0.4	7.2	2.0	7.3	2.3	7.3	**7.5**	1.0	12.4	1.5	12.6	1.5	12.9	**14.1**

This table shows the efficiencies for each (combination of) variance reduction technique, with "A" denoting antithetic sampling, "C" denoting control variates and "I" denoting importance sampling, respectively. For each option, the highest efficiencies for call and symmetric put option are indicated in boldface.



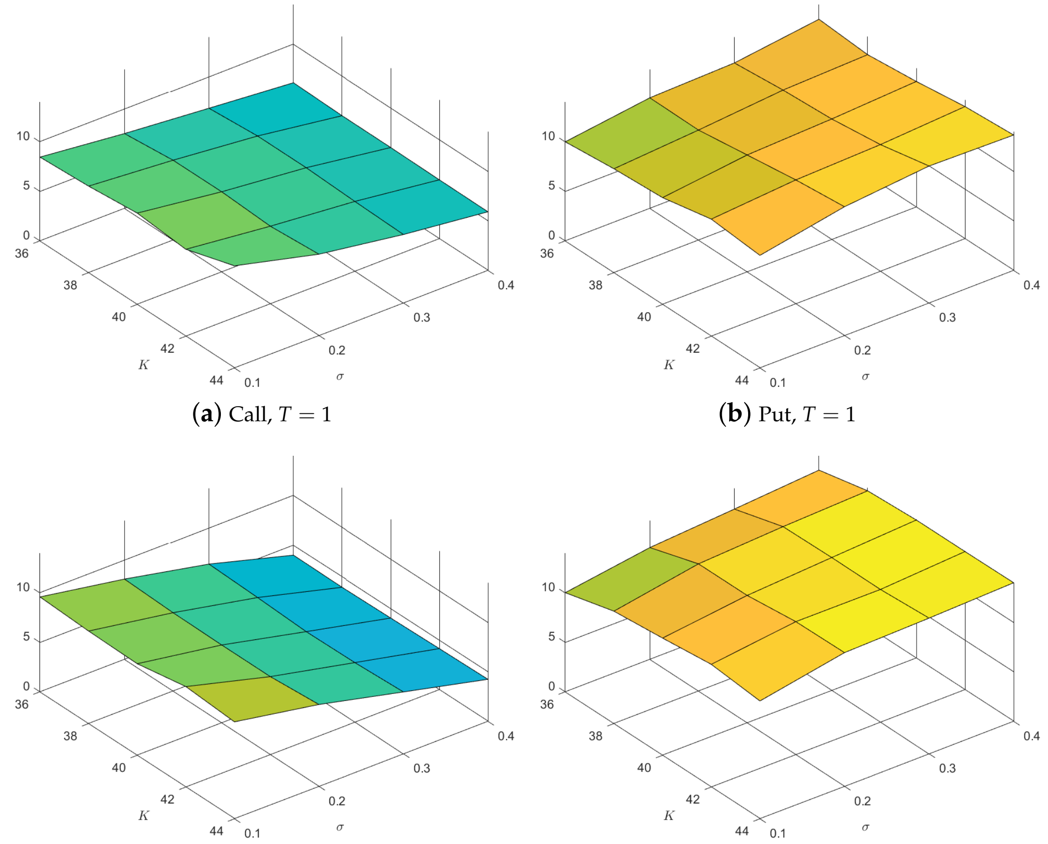
(**a**) Call, $T = 1$ (**b**) Put, $T = 1$

(**c**) Call, $T = 2$ (**d**) Put, $T = 2$

Figure 2. Efficiency of pricing with antithetic variates. Results are calculated from $M = 1000$ replications of LSM and A-LSM configurations with $N = 100{,}000$ paths and a cubic approximation ($L = 3$) of the decision rule. The left column illustrates efficiencies for call options, and the right column for symmetric put options. The top and bottom rows present results for one- and two-year options, respectively.

3.2. *Variance Reduction and Symmetric Pricing*

While the efficiency gains from variance reduction techniques are different for call and (symmetric) put options, once a method that involves control variates is considered Table 1 shows that a given technique (or combination of techniques) works better for put options than for the corresponding call options. For example, when using control variates the efficiency factor is between 5 and 10 for call options, while for the corresponding puts, the efficiency is larger, and nearly always between 10 and 15. This performance is highlighted in Figure 3. The relative improvement is largest when combining all three variance reduction techniques which leads to a $(14.1 - 7.5)/7.5 = 88.0\%$ larger variance reduction for the put than for the call options. Since the variance is always lower for the symmetric put than for the corresponding call option in Figure 1 it follows that the variance reduction that can be achieved for call options could be improved upon by considering variance reduction techniques together with symmetrical pricing.



(**a**) Call, $T = 1$ (**b**) Put, $T = 1$

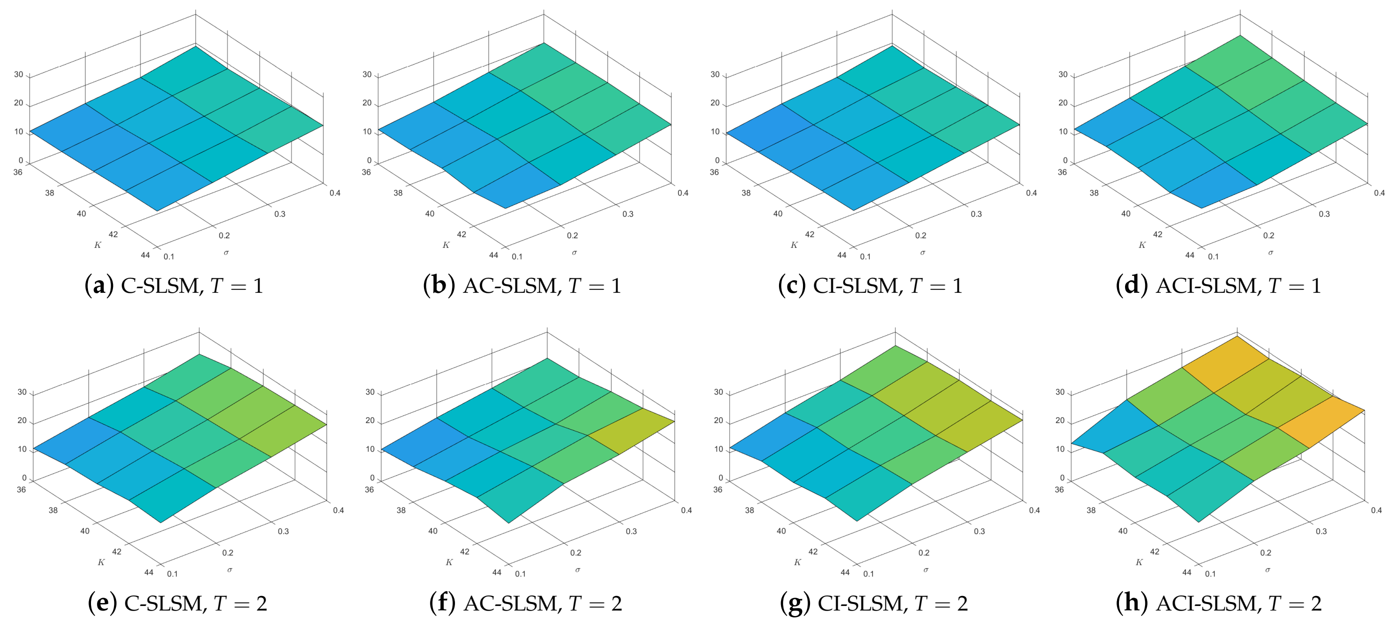
(**c**) Call,$T = 2$ (**d**) Put, $T = 2$

Figure 3. Efficiency of pricing with control variates. Results are calculated from $M = 1000$ replications of LSM and C-LSM configurations with $N = 100{,}000$ paths and a cubic approximation ($L = 3$) of the decision rule. The left column illustrates efficiencies for call options, and the right column for symmetric put options. The top and bottom rows present results for one- and two-year options, respectively.

To analyse the joint effect of variance reduction and symmetric pricing, we now consider the *total* efficiency for call options of a joint implementation of symmetric pricing and various (combinations of) variance reduction techniques. We define the total efficiency as a slight modification of the efficiency measure in Equation (10) given by

$$\text{Tot Eff}_{VR} = \sqrt{\frac{\text{Var}\left[\bar{V}_{call}^{(N)}\right]}{\text{Var}\left[\bar{V}_{put,VR}^{(N)}\right]}} - 1, \tag{11}$$

where $\bar{V}_{call}^{(N)}$ is a crude Monte Carlo estimator for a call option using a sample of N paths, and $\bar{V}_{put,VR}^{(N)}$ is the symmetric put estimator using the same sample supplemented with a (combination of) variance reduction technique(s). Again, the variances are calculated across the $M = 1000$ replications. Figure 4 depicts the total efficiency for all the possible combinations of variance reduction techniques considered that involve control variates for the 1-year and 2-year call options in the top row and bottom row, respectively. This figure can be compared to Figures A3 and A4 in Appendix B which shows the efficiency values for call and (symmetric) puts, respectively.



(**a**) C-SLSM, $T = 1$ (**b**) AC-SLSM, $T = 1$ (**c**) CI-SLSM, $T = 1$ (**d**) ACI-SLSM, $T = 1$

(**e**) C-SLSM, $T = 2$ (**f**) AC-SLSM, $T = 2$ (**g**) CI-SLSM, $T = 2$ (**h**) ACI-SLSM, $T = 2$

Figure 4. Total efficiency of symmetric pricing with (combinations of) control variates. Results are calculated from $M = 1000$ replications of LSM configurations with $N = 100{,}000$ paths and a cubic approximation ($L = 3$) of the decision rule. The first column illustrates total efficiencies of C-SLSM prices, the second column for AC-SLSM, the third column for CI-SLSM, and the fourth column for ACI-SLSM. The top and bottom rows present results for one- and two-year options, respectively.

Figure 4 clearly suggests that symmetric pricing is very effective when implemented with control variates. Indeed, for high volatility options, symmetric pricing reduces the variance of the LSM estimator of the call option price by a factor of more than 20 when combining all three variance reduction techniques. These numbers are much higher than when implementing variance reduction without symmetry shown in Table 1 and plotted in Figure A3 in Appendix B. Although symmetric pricing implemented with antithetic and importance sampling alone is also more effective than in a vanilla LSM configuration, the total gain in efficiency is much lower. In fact, our results show that when pricing call options, combining variance reduction techniques with symmetric pricing always leads to more precise estimates. Thus, when considering variance reduction techniques one should also consider using symmetric pricing, and conversely when considering pricing call options using symmetry one should also consider implementing variance reduction techniques for the put option.

In Table 2 we report the results when using an out-of-sample pricing approach instead. Here the optimal stopping time is estimated from the standard LSM method, i.e., without variance reduction, and the same strategy is used for all the out-of-sample pricing with (combinations of) variance reduction techniques. The first thing to notice is that the results are very similar to those reported in Table 1 demonstrating that our findings are not particular to using in-sample pricing. Specifically, the effect of variance reduction techniques differ between call and put options, but control variates should always be used.

Second, Table 2 also shows that when a control variate is used, either by itself or in combination with other techniques, variance reduction is much more efficient for put options than for call options. One minor difference is that for out-of-sample pricing, ACI-LSM improves somewhat less on the other strategies on average, compared to when using in-sample pricing. However, on average, combining all three variance reduction techniques still offers the highest efficiency, particularly when applied to put options. As a result, combining variance reduction using control variates with symmetric pricing offers huge benefits also when doing out-of-sample call option pricing.[4]

Table 2. Efficiency results for out-of-sample pricing.

			Call							Symmetric Put						
T	K	σ	**A**	**C**	**I**	**AC**	**AI**	**CI**	**ACI**	**A**	**C**	**I**	**AC**	**AI**	**CI**	**ACI**
1	36	0.1	0.5	7.6	1.1	7.6	1.1	7.6	**7.7**	0.4	9.2	0.9	9.3	1.0	9.2	**9.4**
1	36	0.2	1.1	7.1	1.4	7.1	1.4	7.1	**7.1**	1.1	10.3	1.0	10.3	1.1	10.4	**10.6**
1	36	0.3	0.9	6.2	1.5	6.2	1.6	6.2	**6.2**	1.4	11.7	1.0	11.7	1.1	11.8	**12.0**
1	36	0.4	0.6	5.6	1.7	5.6	1.8	5.6	**5.6**	1.6	12.3	1.0	12.2	1.1	12.4	**12.6**
1	38	0.1	1.2	8.1	1.3	8.1	1.4	8.1	**8.1**	1.2	9.8	1.1	9.8	1.2	9.8	**10.0**
1	38	0.2	0.9	7.5	1.5	7.5	1.6	7.6	**7.6**	1.4	11.4	1.2	11.5	1.2	11.4	**11.6**
1	38	0.3	0.6	6.6	1.6	6.6	1.7	6.7	**6.7**	1.5	12.0	1.1	12.2	1.2	12.1	**12.3**
1	38	0.4	0.4	5.6	1.8	5.6	1.8	5.6	**5.6**	1.6	12.7	1.0	12.8	1.1	12.7	**12.9**
1	40	0.1	0.5	8.6	1.6	8.6	1.6	8.6	**8.7**	0.8	10.7	1.4	10.8	1.4	10.8	**10.8**
1	40	0.2	0.5	7.8	1.7	7.8	1.7	**7.9**	7.8	0.9	11.9	1.3	**11.9**	1.3	11.9	11.9
1	40	0.3	0.4	7.0	1.8	7.0	1.8	7.0	**7.0**	1.1	12.7	1.2	**12.8**	1.2	12.7	12.8
1	40	0.4	0.3	5.8	1.9	5.8	1.9	5.8	**5.8**	1.3	13.1	1.1	13.2	1.2	13.2	**13.3**
1	42	0.1	0.2	9.3	2.2	9.3	2.1	**9.3**	9.3	0.2	11.3	1.8	11.3	1.8	11.4	**11.4**
1	42	0.2	0.3	8.1	2.0	8.0	2.0	**8.1**	8.1	0.4	11.8	1.4	11.7	1.5	11.9	**11.9**
1	42	0.3	0.3	7.0	2.0	7.0	2.0	**7.0**	7.0	0.7	12.9	1.3	12.8	1.4	12.9	**13.0**
1	42	0.4	0.3	5.9	2.0	5.9	2.0	**5.9**	5.9	0.9	13.4	1.2	13.4	1.2	13.4	**13.5**
1	44	0.1	0.1	10.4	2.9	10.4	2.8	10.5	**10.5**	0.0	**10.8**	2.4	10.8	2.3	10.8	10.8
1	44	0.2	0.2	8.4	2.2	8.4	2.2	8.4	**8.4**	0.2	12.1	1.7	12.1	1.7	12.3	**12.3**
1	44	0.3	0.2	7.3	2.2	7.3	2.1	**7.3**	7.3	0.4	12.7	1.4	12.7	1.5	12.8	**12.8**
1	44	0.4	0.2	6.1	2.1	6.1	2.1	**6.1**	6.1	0.6	13.3	1.3	13.2	1.3	13.3	**13.4**
2	36	0.1	0.7	9.7	1.8	9.8	1.8	9.6	**10.1**	0.7	10.5	1.6	11.0	1.5	10.4	**11.2**
2	36	0.2	1.1	8.4	2.3	8.5	2.2	8.6	**8.7**	1.2	12.2	1.5	12.1	1.5	12.4	**13.3**
2	36	0.3	0.7	6.5	2.3	6.6	2.3	6.6	**6.6**	1.4	12.8	1.4	12.7	1.4	13.1	**14.0**
2	36	0.4	0.5	4.5	2.2	4.5	2.3	4.5	**4.5**	1.4	13.0	1.3	12.8	1.2	13.4	**14.4**
2	38	0.1	1.3	10.1	2.2	10.2	2.1	10.5	**10.7**	1.3	12.2	1.7	12.2	1.6	12.4	**13.0**
2	38	0.2	0.7	8.2	2.2	8.4	2.2	8.4	**8.5**	1.4	12.4	1.5	12.7	1.5	13.1	**13.7**
2	38	0.3	0.6	6.5	2.4	6.6	2.4	6.6	**6.6**	1.5	12.7	1.4	12.9	1.4	13.3	**14.0**
2	38	0.4	0.4	4.2	2.1	4.2	2.1	4.2	**4.2**	1.5	13.4	1.3	13.6	1.3	14.0	**14.9**
2	40	0.1	0.6	10.1	2.2	10.2	2.1	10.3	**10.5**	0.9	12.4	1.7	12.9	1.7	13.3	**13.6**
2	40	0.2	0.5	8.3	2.3	8.3	2.2	8.4	**8.5**	1.1	13.0	1.6	13.6	1.5	13.9	**14.4**
2	40	0.3	0.4	6.2	2.3	6.2	2.3	6.2	**6.2**	1.3	13.5	1.5	14.0	1.4	14.5	**15.0**
2	40	0.4	0.3	4.1	2.1	4.1	2.1	4.2	**4.2**	1.6	13.6	1.4	13.9	1.3	14.6	**15.3**
2	42	0.1	0.3	10.2	2.3	10.2	2.3	10.3	**10.4**	0.4	11.9	1.9	11.9	1.9	12.4	**12.5**
2	42	0.2	0.4	8.2	2.4	8.1	2.4	8.2	**8.3**	0.7	13.4	1.7	13.8	1.7	14.3	**14.6**
2	42	0.3	0.3	6.2	2.3	6.1	2.3	6.2	**6.3**	1.0	14.1	1.6	14.7	1.5	15.2	**15.7**
2	42	0.4	0.3	4.2	2.1	4.2	2.2	4.2	**4.3**	1.4	14.5	1.5	15.3	1.3	15.7	**16.6**
2	44	0.1	0.1	9.5	2.5	9.7	2.7	9.7	**9.8**	0.2	12.2	2.1	12.1	2.2	12.6	**12.6**
2	44	0.2	0.3	7.9	2.4	7.9	2.4	8.0	**8.0**	0.5	12.9	1.8	13.0	1.8	13.6	**13.7**
2	44	0.3	0.3	5.9	2.3	5.9	2.4	5.9	**5.9**	0.8	14.2	1.7	14.7	1.6	15.3	**15.7**
2	44	0.4	0.3	4.4	2.2	4.4	2.2	4.4	**4.4**	1.2	14.7	1.5	15.4	1.4	15.9	**16.5**
Average			0.5	7.2	2.0	7.2	2.0	7.3	**7.3**	1.0	12.4	1.4	12.5	1.4	12.8	**13.1**

This table shows the efficiencies for each (combination of) variance reduction technique, with "A" denoting antithetic sampling, "C" denoting control variates and "I" denoting importance sampling, respectively. For each option, the highest efficiencies for call and symmetric put option are indicated in boldface.

3.3. Discussion

In the following, we provide some further intuition for why some variance reduction techniques are more efficient than others. Before doing so, it is important to mention that these comparisons only concern vanilla put and call options, and our conclusions should therefore be interpreted in the specific context of pricing American options with simple payoff functions. There is generally no best-performing variance reduction technique as efficiencies are largely context-specific. A good rule of thumb, as pointed out by Glasserman (2003), is that the more information leveraged about the option properties, the larger the reduction in variance. In practise, however, leveraging this information, and optimizing the variance reduction tools for a specific problem can be so computationally taxing that it would be ill-advised to do so, even if the efficiency gains are large. The success of variance reduction techniques therefore rests on striking a balance between efficiency, and ease of implementation.

On one end of the spectrum, the simplest tool is antithetic sampling, as it requires no knowledge of the option whatsoever, and, as the rule suggests, generally offers lower efficiencies than the two others. Our results show that this intuition applies to our setting in which antithetic sampling has the lowest efficiencies. More generally, we conjecture that the efficiency of antithetic sampling in the context of LSM is subdued in comparison to an European option estimator. Indeed, the efficiency gains permitted by antithetic sampling are larger as the correlation between antithetically simulated option payoffs becomes negative. Supposing that a particular pair of antithetic paths $\{S_{n,j}, \tilde{S}_{n,h}, j = 0, \ldots, J\}$ is obtained from the normal increments $\{Z_n, -Z_n\}$, respectively, the LSM will estimate stopping times $\{k_n, \tilde{k}_n\}$ corresponding to each path in the antithetic pair, and generate a pair of discounted payoffs $\{e^{-r\tau_{k_n}} h(S_{n,k_n}), e^{-r\tau_{\tilde{k}_n}} h(\tilde{S}_{n,\tilde{k}_n})\}$. The correlation between antithetic payoffs is necessarily negative when stopping times coincide (i.e., $k_n = \tilde{k}_n$), as $h(\cdot)$ is a monotone increasing function in the asset price, and antithetic pairs of asset prices are negatively correlated at a given time. For European options, all exercise times occur at maturity (i.e., $k_n = \tilde{k}_n = J$), and it immediately follows that antithetic sampling necessarily provides an improvement over the standard Monte Carlo approach. This is not generally the case for American options, though, because the antithetically estimated stopping times rarely coincide. Furthermore, if both paths within an antithetic pair are exercised, this will occur at different times and the discounted payoffs will be more positively correlated than pairs of paths having exactly one positive payoff. The resulting effect is a weaker overall correlation, and a subdued efficiency of antithetic sampling.

On the other end of the spectrum, control variates are the most efficient across all option properties, and, as expected, require extensive knowledge about the characteristics of the option. Our experiments consider an ideal case in which the option that serves as a control variate can be valued easily, and replicates the American option payoff exactly at maturity.[5] These ideal control variates are nearly perfectly correlated with the option values, yielding very large efficiencies across all levels of maturity, asset volatility, and moneyness. In essence, the efficiency of control variates rests on the knowledge of the price of "nearby" options. However, it is worth noting that the implementation of control variates is more challenging and more computationally taxing than the other variance reduction techniques we consider. In particular, at every time step, the optimal control variates method proposed by Rasmussen (2005) requires one to compute a cross-section of European option prices. Moreover, the large efficiency gains permitted by control variates largely depend on the optimization of control parameters, which requires one to perform three additional regressions at every step to estimate the correlation between control variates and estimated payoffs.

In the middle of the spectrum, we can place our implementation of importance sampling with a uniform change of drift. This simple parameterization of the importance probability measure uses only limited knowledge of the option characteristics and leads therefore to modest efficiency gains. This variance reduction technique depends on the optimization of the drift parameter, and an ill-chosen drift may very well lead to the estimator having larger variance. If we decided to optimize the importance measure over a larger parametric space (e.g., if we decided to optimize for the drift and the volatility jointly), further gains in efficiency may be obtained, and perhaps an optimal implementation of importance sampling would outperform control variates. However, as discussed in Appendix A.3, the optimization of a drift term is already challenging task, and furthermore, as detailed in Appendix A.2, the optimization is a delicate process when importance sampling is combined with other variance reduction techniques. That being said, the optimization of importance probability measures with more general parameterizations is still an interesting problem for American option pricing, but one we leave for future research.[6]

4. Conclusions

Efficient pricing of American options remains an active area of research. Among the many numerical methods that exist, regression based simulation methods have become popular due to their flexibility. However, as is the case with other Monte Carlo based methods they are numerically costly and reducing the variance of such methods is therefore important. This paper conducts a thorough study of the effect of using various standard variance reduction methods and combinations thereof for a sample of call and (symmetric) put options whose solutions coincide.

Our results first show that the efficiency gains from variance reduction techniques are different for call and put options, and methods that involve control variates work the best. Moreover, any combination of variance reduction methods that include control variates works better for the (symmetric) put option than for the corresponding call option. Finally, since symmetric pricing already offers some variance reduction, our results show that one may obtain reductions in the standard deviation by a factor of more than 20 of a call option when combining control variates with symmetric pricing.

The improvement is largest for long maturity options on high volatility assets. Our suggested approach is therefore of particular relevance to the problem of valuing real options, as they typically take the form of deep OTM, long maturity call options on volatile assets. In various sectors, the valuation of real options is of paramount importance for the financial flexibility of the firm, because development projects can be extremely costly, have very long maturities, and the resulting cashflows are difficult to predict. Indeed, in the energy sector, for instance, the proper management of such options adds tremendous value to the firms operations. Because our proposed method is limited to cases where the symmetry property holds, the applicability of the symmetric pricing to real option pricing is left for future research. Extensions to multiple exercise options and alternative stochastic processes where symmetry applies is also left for future work.

Although our general recommendation is to use control variates and symmetric pricing when pricing call options, there is some potential for improving efficiency further for high volatility options by combining this with antithetic and importance sampling. In this case there does not appear to be any offsetting effects of the variance reduction techniques, though implementing the combined techniques is much more challenging. We also leave the question of whether our recommended approach is "optimal" when considering the trade off between computational costs and precision for future research.

Author Contributions: Conceptualization, F.-M.B., R.M.R. and L.S.; data curation, F.-M.B. and R.M.R.; formal analysis, F.-M.B. and R.M.R.; funding acquisition, F.-M.B. and R.M.R.; project administration, L.S.; supervision, L.S. All authors have read and agreed to the published version of the manuscript.

Funding: F.-M.B. thanks the University of Western Ontario and OGS for financial support. R.M.R. thanks NSERC and Wilfrid Laurier University for financial support. L.S. thanks NSERC for financial support. This work was supported by the Research Support Fund.

Data Availability Statement: Data sharing is not applicable to this article.

Acknowledgments: The authors would like to thank three anonymous referees for valuable comments and SHARCNET for providing computational resources.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Variance Reduction for American Option Prices

This appendix first introduces the three variance reduction methods we consider in this paper. Then it provides a discussion of some numerical issues that arises when implementing these techniques. Finally, it provides some evidence on the challenge that arises when choosing the optimal importance sampling measure when combining this variance reduction technique with other techniques, in particular.

Appendix A.1. Variance Reduction Techniques

Let $\hat{g}(Z_n)$ be the path-n discounted payoff of an American option resulting from an approximate solution to the dynamic program. The sample size-N crude estimator of the American option is then

$$\bar{V}^{(N)} = N^{-1} \sum_{n=1}^{N} \hat{g}(Z_n), \tag{A1}$$

where the sample paths are generated with independent random normal increments $Z_n = \{z_{n,1}, \ldots, z_{n,J}\}$, where $z_{n,j} \overset{IID}{\sim} \mathcal{N}(0,1) : n = 1, \ldots, N; j = 1, \ldots, J$. Antithetic sampling, control variates, and importance sampling are widely used to improve on such estimators. We restrict our focus to these three techniques because they require no additional simulated paths. In this way, the estimated exercise strategies are derived from similar samples, allowing for a fair comparison of the efficiency of each technique. Other variance reduction techniques that are not examined here include latin hypercube sampling (Broadie et al. 1997), local policy enhancement (Broadie and Cao 2008), and low discrepancy sequences (Lemieux and La 2005). Initial state dispersion (Rasmussen 2005) may also indirectly function as a variance reduction technique.

The idea behind antithetic sampling is that one can construct a new estimator by combining $\hat{g}(Z_n)$ and $\hat{g}(-Z_n)$, the latter of which is based on the antithetic sample of random variates. For a fair assessment of the efficiency of antithetic sampling, we say that an antithetic estimator with N sample paths is the average of $N/2$ antithetic pairs of regular estimators such that

$$\bar{V}_A^{(N)} = N^{-1} \sum_{n=1}^{\frac{N}{2}} (\hat{g}(Z_n) + \hat{g}(-Z_n)). \tag{A2}$$

Antithetic sampling was used with the LSM method in the original paper by Longstaff and Schwartz (2001).

The idea behind control variates is instead to correct (at least partially) for the sample variance by using an available unbiased estimator $\hat{f}_N = N^{-1} \sum_{n=1}^{N} \hat{f}(Z_n)$ of a known quantity f. Specifically, a controlled estimator is obtained from

$$\bar{V}_C^{(N)} = N^{-1} \sum_{n=1}^{N} \Big(\hat{g}(Z_n) + \theta \Big(\hat{f}(Z_n) - f \Big) \Big), \tag{A3}$$

where θ is the *control coefficient*. A natural candidate for a control variate is the price of the corresponding European price or the price of a derivative when using simpler dynamics. Control variates was used with the LSM method as early as in Tian and Burrage (2002).

The idea behind importance sampling is to adjust the paths in such a way that more of them have non-zero payoffs and therefore contain information about the value of the option. In the simplest possible implementation a drift term is added to the normal increments by posing $\tilde{Z}_n \equiv \{z_{n,1} + \sqrt{\Delta t}\lambda, \ldots, z_{n,J} + \sqrt{\Delta t}\lambda\}$. Parameterizing the drifted probability measure as $\mathbb{P}^\lambda$, the importance sampling estimator is then given by

$$\bar{V}_I^{(N)} = N^{-1} \sum_{n=1}^{N} \hat{g}(\tilde{Z}_n) \frac{d\mathbb{P}}{d\mathbb{P}^\lambda}(S_{n,\tau(n)}), \tag{A4}$$

where the last term is the likelihood ratio. Thus, pathwise discounted cashflows are first computed with paths simulated under $\mathbb{P}^\lambda$ and subsequently multiplied with the corresponding likelihood ratio. Importance sampling was used with the LSM method as early as in Moreni (2003).

Appendix A.2. Implementation Issues

Having provided intuition for the different variance reduction techniques used in this paper, we now highlight some numerical issues which may arise when implementing (combinations of) variance reduction techniques in a dynamic programming context. We call attention to three issues: (i) the estimation of an exercise rule, (ii) the selection of optimal control variates and (iii) the (automated) choice of an importance measure used for importance sampling.

First, when a variance reduction technique modifies the simulated sample paths, as is the case with antithetic and importance sampling, it impacts the exercise rule approximation and possibly the bias of an LSM estimator, whereas in a simple Monte Carlo integration problem (e.g., with a known exercise strategy), the bias remains unchanged. The approximation of an exercise rule is a source of LSM estimator bias and variance, and since OLS assumptions of regressor independence are clearly violated with antithetic sampling, this impacts the estimated continuation values and hence exercise decisions. Antithetic sampling does not significantly change the bias when the sample size is sufficiently large (100,000 simulated paths), but for smaller sample sizes inadequate OLS estimates may lead to a poor exercise strategy, resulting in inordinately large negative bias. The problem when using importance sampling is that the optimal stopping time strategy determined under the nominal measure is applied to paths simulated under the importance density which may lead to negatively biased results. Note that in our simple setting a single simulation procedure is sufficient to obtain sample paths under both measures, because cross-sections of asset paths are co-linear, however importance sampling may lead to higher memory requirements if two sets of simulated paths need to be stored.

Second, when implementing control variates a proper control has to be chosen. An obvious control to use is the equivalent European price. However, as demonstrated by Rasmussen (2005) a more effective choice of control variates is the European option price sampled pathwise at the exercise time of the American option. Specifically, if the path-n LSM-estimated stopping time is $\tau(n)$, the path-n control variate used is the discounted value of a European option with maturity $T - \tau(n)$ and initial asset price $S_{\tau(n)}$. Thus, we set $\hat{f}(Z_n) = e^{-r\tau(n)}\mathbb{E}[e^{-r(T-\tau(n))}h(S_T) \mid \mathcal{G}_{\tau(n)}]$ and $f = \mathbb{E}[e^{-rT}h(S_T) \mid \mathcal{G}_0]$, which is the value of a European option with maturity T and initial asset price S_0. This technique was first discussed in the context of a stochastic mesh approach (Broadie and Glasserman 2004) and allows remarkable improvements for Monte Carlo variance reduction. Note that control variates do not affect the determination of an exercise rule, as they are only computed at the end of the backward recursion, once the stopping times are estimated.

Third, with no prior knowledge of the true option price, choosing the optimal λ is not feasible and instead we must rely on an approximation. A common approach is to determine the optimal change of measure for a European option and apply it to the corresponding American option, as discussed by Moreni (2003). The benefit is that optimization procedures for European options can be carried out very quickly with Robbins-Monro algorithms (see for instance Arouna 2004), with virtually no overhead. In our experiments, we follow Lemieux and La (2005) and implement the saddle point approximation approach of Glasserman et al. (1999). With this approach, the drift λ is easily computed numerically as

$$\lambda = \begin{cases} \underset{x:x>\xi}{\operatorname{argmax}} \log\left[S_0 \exp\left(\left(r - q - \frac{\sigma^2}{2}\right)T + \sigma\sqrt{T}x\right) - K\right] - \frac{x^2}{2} & \text{for calls} \\ \underset{x:x<\xi}{\operatorname{argmax}} \log\left[K - S_0 \exp\left(\left(r - q - \frac{\sigma^2}{2}\right)T + \sigma\sqrt{T}x\right)\right] - \frac{x^2}{2} & \text{for puts,} \end{cases} \tag{A5}$$

where $\xi = \frac{\log\left(\frac{K}{S_0}\right) - \left(r - q - \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}$. Other methods proposed to optimize the drift include Su and Fu (2000); Vázquez-Abad and Dufresne (1998), and Morales (2006) who employ gradient-based minimization methods to directly optimize the drift. Because the variance of importance sampling estimators is convex with respect to the drift parameter under suitable conditions, numerical convergence is achieved fairly quickly and reliable drift estimates can be obtained. However, as we show in the following section this approach may not be appropriate when importance sampling is combined with other variance reduction techniques.

Appendix A.3. Selecting the Optimal Importance Density

The selection of an importance density is a delicate problem where a simple suboptimal approach is often preferred to an optimal one requiring a hefty implementation effort. In our implementation, we simply optimize the importance density for the European case with a uniform drift of the normal increments, keeping the variance unchanged. An important question is then: how much efficiency are we giving up with this suboptimal European solution in comparison to the optimal density for an American option? To answer this question, we study the variance of I-LSM and AI-LSM estimators with respect to the drift of the Brownian increments. As an example, we consider a deep OTM, long maturity, high volatility American call option with $S_0 = 40$, $K = 44$, $\sigma = 0.4$, $r = q = 0.06$, $T = 2$, $J = 50T$. For a range of values of λ, $M = 50$ replications of the LSM are implemented each with $N = 100{,}000$ paths and a cubic approximation of the continuation value (i.e., $L = 3$). Figures A1 and A2 shows the relative variance as a function of the drift and the estimated prices with 95% confidence intervals.

From Figure A1 the main takeaway is that for stand-alone importance sampling, the Glasserman et al. (1999) (GHS) approximation of the optimal European drift is indeed very close to the American optimum. Moreover, the GHS drift yields gains in efficiency that are similar to what we would otherwise obtain with the true optimum. In such cases, we confirm that it is not worthwhile to undertake computationally taxing gradient descent schemes to reach the true optimum. From Figure A2, however, the main takeaway is that the convexity of the variance function is not preserved when importance sampling is coupled with antithetic sampling. This is very important for practitioners who intend to compute a gradient-based type of solution, because a preliminary search of the optimization surface is warranted, thereby adding to the optimizing burden. From Figure A2 we see that the GHS drift performs slightly worse than the true optimum, but the variance reduction is still substantial. Note that the optimal drift value will also change when importance sampling is combined with control variates.

Finally, one should also note that gradient-based solutions are subject to a number of numerical defects. First, the optimum is sensitive to the number of simulation paths, particularly for small samples sizes. A large number of simulation paths is then required to obtain stable solutions that match those in the literature. Second, numerical convergence issues can arise if the gradient descent does not lead to solutions in a certain neighborhood of the optimum. For instance, if the drift leads all paths to be OTM, the pathwise cashflows will all be zero and the variance is zero. Similarly, if the drift leads all paths to be deep ITM, the likelihood ratio will converge to zero, and the weighted pathwise payoffs will be infinitesimally small. Of course, these issues are easily identifiable as they lead to inordinately biased prices. Still, meticulous care is advised with such delicate algorithms (Moreni 2003), and they might not be useful for valuating a large panel of options in an automated procedure. All things considered, we use the approximate optimal European drift of Glasserman et al. (1999) due to its ease of implementation and realized efficiency.

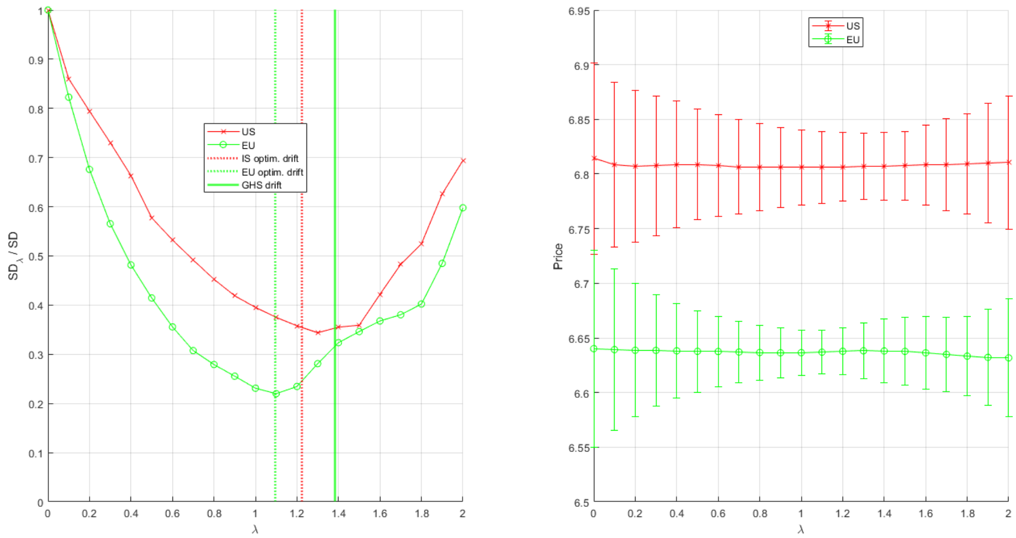


Figure A1. Variance and price estimates of the I-LSM estimator as a function of λ. Results for the American call option appear in red, and the results for European call price estimators with importance sampling appear in green. For each value of λ, $M = 50$ replications of the estimators are implemented with $N = 100{,}000$ paths. LSM and I-LSM price estimates use a cubic approximation of the continuation value (i.e., $L = 3$). The left panel shows the ratio of estimator variances compared to an estimator under the nominal probability measure (i.e., $\lambda = 0$). The optimal values of λ selected by gradient-descent, and that selected by the GHS approach appear as dotted, and full vertical lines, respectively. The right panel shows price estimate with 95% confidence intervals.

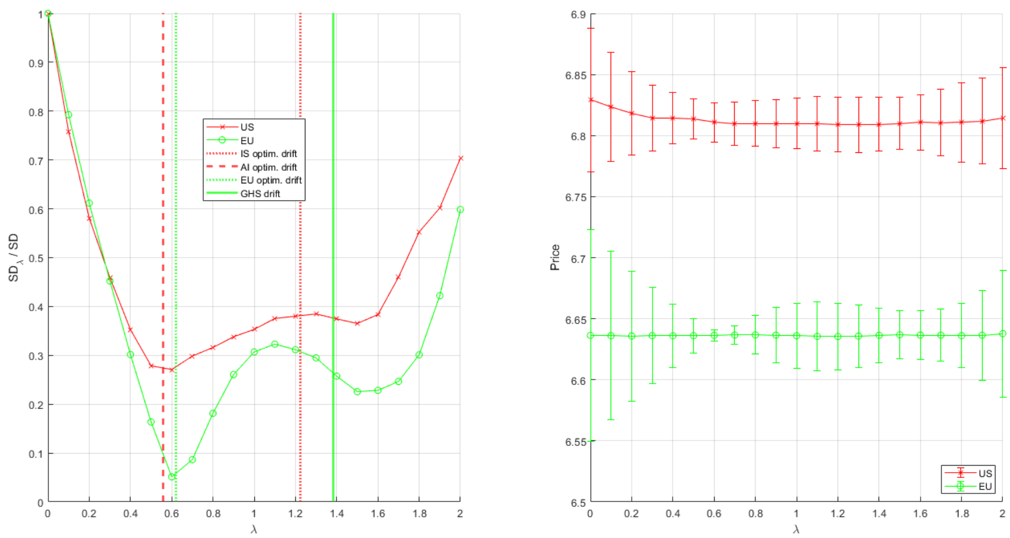
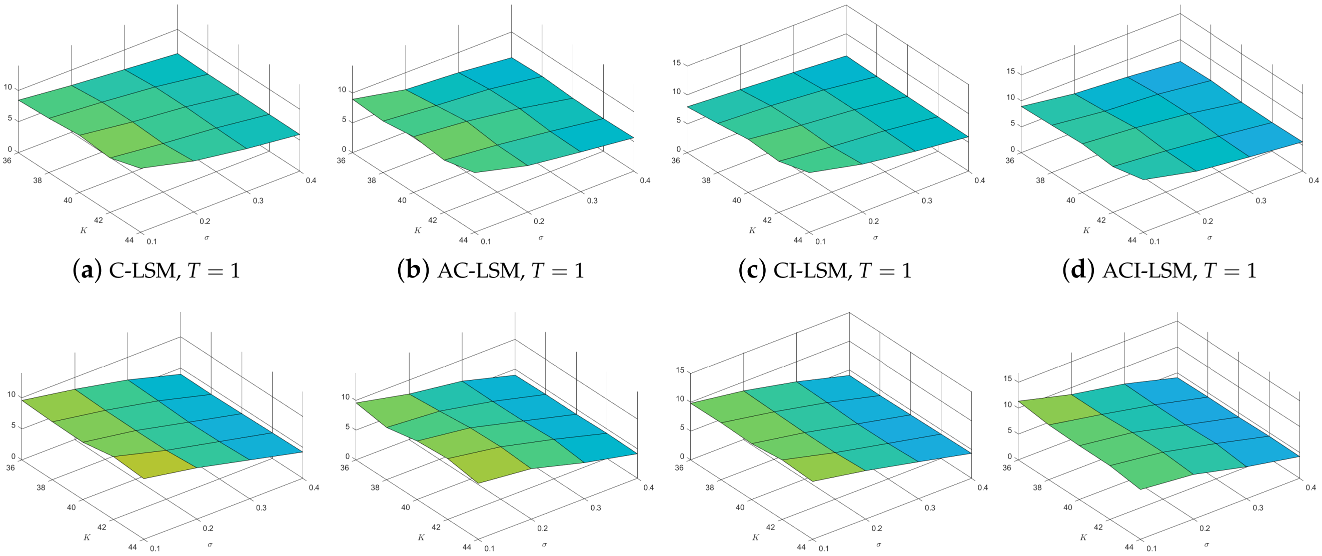


Figure A2. Variance and price estimates of the AI-LSM estimator as a function of λ. Results for the American call option appear in red, and the results for European call price estimators that combine antithetic variates and importance sampling appear in green. For each value of λ, $M = 50$ replications of the estimators are implemented with $N = 100{,}000$ paths. LSM and AI-LSM price estimates use a cubic approximation of the continuation value (i.e., $L = 3$). The left panel shows the ratio of estimator variances compared to an estimator under the nominal probability measure (i.e., $\lambda = 0$). The optimal values of λ selected by gradient-descent for AI- and I-LSM estimators, and that selected by the GHS approach appear as dashed, dotted, and full vertical lines, respectively. The right panel shows price estimate with 95% confidence intervals.

Appendix B. Additional Figures

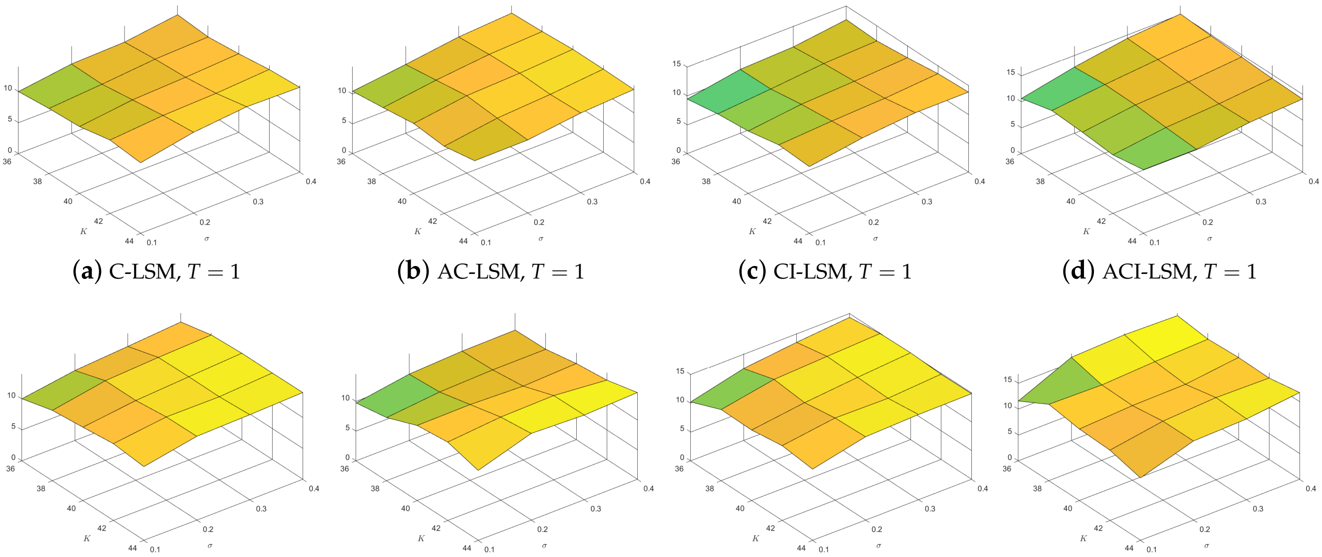
This appendix contains additional figures complementing those presented in the body of the paper.



(**a**) C-LSM, $T = 1$ (**b**) AC-LSM, $T = 1$ (**c**) CI-LSM, $T = 1$ (**d**) ACI-LSM, $T = 1$

(**e**) C-LSM, $T = 2$ (**f**) AC-LSM, $T = 2$ (**g**) CI-LSM, $T = 2$ (**h**) ACI-SLSM, $T = 2$

Figure A3. Efficiency for call option pricing with (combinations of) control variates. Results are calculated from $M = 1000$ replications of LSM configurations with $N = 100{,}000$ paths and a cubic approximation ($L = 3$) of the decision rule. The first column illustrates efficiencies of C-LSM call prices, the second column for AC-LSM, the third column for CI-LSM, and the fourth column for ACI-LSM. The top and bottom rows present results for one- and two-year options, respectively.



(**a**) C-LSM, $T = 1$ (**b**) AC-LSM, $T = 1$ (**c**) CI-LSM, $T = 1$ (**d**) ACI-LSM, $T = 1$

(**e**) C-LSM, $T = 2$ (**f**) AC-LSM, $T = 2$ (**g**) CI-LSM, $T = 2$ (**h**) ACI-LSM, $T = 2$

Figure A4. Efficiency for put option pricing with (combinations of) control variates. Results are calculated from $M = 1000$ replications of LSM configurations with $N = 100{,}000$ paths and a cubic approximation ($L = 3$) of the decision rule. The first column illustrates efficiencies of C-LSM put prices, the second column for AC-LSM, the third column for CI-LSM, and the fourth column for ACI-LSM. The top and bottom rows present results for one- and two-year options, respectively.

Notes

1 For additional references and examples see Chapter 8 of Glasserman (2003).

2 In our implementation we simply set $\psi_\ell(S_{n,j}) = (S_{n,j}/K)^\ell$ such that the continuation value is the fitted value of a polynomial regression of order $L < \infty$. This approach has been shown to be reliable with the LSM, though other orthogonal bases like Laguerre, Legendre, Hermite, or Chebyshev polynomials may be considered as well. For more details about orthogonal bases, refer to Abramowitz and Stegun (1948).

3 Appendix A discusses the variance reduction techniques and their implementation for American options.

4 The out-of-sample plots corresponding to Figure 4 that shows this are available on request.

5 When the price of a European option replicating the payoff of the American option at maturity is not readily available, one has to resort to using different European options with simpler characteristics. For instance, there is no closed-form formula for the price of an arithmetic Asian option, so the price of a geometric Asian option may be used as a control variate. Another case is if the computation of a European option prices is not feasible for a given underlying diffusion process, where the European price assuming a different process can be used. Although in both cases the European option does not replicate the American option's payoff at maturity, their values are correlated enough to serve as adequate control variates.

6 Other interesting sampling strategies for European option pricing could potentially be extended to the LSM algorithm. Chapter 4 of Glasserman (2003) presents several cases of European-style derivatives with rare and path-dependent payoffs where the application of importance sampling and stratification techniques prove extremely efficient. For instance, in the case of a deep OTM knock-in option, the exponential change of measure of the option can be dynamically adjusted such that asset prices are first directed toward the barrier, and once the barrier has been hit, a new change of measure directs asset prices toward the money.

References

Abramowitz, Milton, and Irene A. Stegun. 1948. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Washington, DC: US Government Printing Office, vol. 55.

Arouna, Bouhari. 2004. Robbins-Monro Algorithms and Variance Reduction in Finance. *Journal of Computational Finance* 7: 35–62. [CrossRef]

Barraquand, Jérôme, and Didier Martineau. 1995. Numerical Valuation of High Dimensional Multivariate American Securities. *Journal of Financial and Quantitative Analysis* 30: 383–405. [CrossRef]

Bolia, Nomesh, Sandeep Juneja, and Paul Glasserman. 2004. Function-Approximation-Based Importance Sampling for Pricing American Options. Paper presented at the 2004 Winter Simulation Conference, Washington, DC, USA, December 5–8. Proceedings (IEEE Cat. No.04CH37614C). vol. 1, pp. 604–11.

Brealey, Richard A., Stewart C. Myers, Franklin Allen, and Pitabas Mohanty. 2018. *Principles of Corporate Finance, 12/e*. New York: McGraw-Hill Education, vol. 12.

Broadie, Mark, and Menghui Cao. 2008. Improved Lower and Upper Bound Algorithms for Pricing American Options by Simulation. *Quantitative Finance* 8: 845–61. [CrossRef]

Broadie, Mark, and Paul Glasserman. 2004. A Stochastic Mesh Method for Pricing High-Dimensional American Options. *Journal of Computational Finance* 7: 35–72. [CrossRef]

Broadie, Mark, Paul Glasserman, and Gautam Jain. 1997. Enhanced Monte Carlo Estimates for American Option Prices. *Journal of Derivatives* 5: 25–44. [CrossRef]

Detemple, Jérôme. 2001. American Options: Symmetry Properties. In *Option Pricing, Interest Rates and Risk Management*. Edited by Elyès Jouini, Jakša Cvitanić and Marek Musiela. Cambridge: Cambridge University Press.

Duan, Jin-Chuan, and Jean-Guy Simonato. 1998. Empirical Martingale Simulation for Asset Prices. *Management Science* 44: 1218–33. [CrossRef]

Glasserman, Paul. 2003. *Monte Carlo Methods in Financial Engineering*. New York: Springer Science & Business Media, vol. 53.

Glasserman, Paul, Philip Heidelberger, and Perwez Shahabuddin. 1999. Asymptotically Optimal Importance Sampling and Stratification for Pricing Path-Dependent Options. *Mathematical Finance* 9: 117–52. [CrossRef]

Juneja, Sandeep, and Himanshu Kalra. 2009. Variance Reduction Techniques for Pricing American Options Using Function Approximations. *Journal of Computational Finance* 12: 79. [CrossRef]

Lemieux, Christiane, and Jennie La. 2005. A Study of Variance Reduction Techniques for American Option Pricing. Paper presented at the 2005 Winter Simulation Conference, Orlando, FL, USA, December 4–7. Proceedings (IEEE Cat. No.05CH37732C). p. 8.

Longstaff, Francis A., and Eduardo S. Schwartz. 2001. Valuing American Options by Simulation: A Simple Least-Squares Approach. *The Review of Financial Studies* 14: 113–47. [CrossRef]

McDonald, Robert L., and Mark D. Schroder. 1998. A Parity Result for American Options. *Journal of Computational Finance* 1: 5–13. [CrossRef]

Morales, Manuel. 2006. *Implementing Importance Sampling in the Least-Squares Monte Carlo Approach for American Options*. Montreal: University of Montreal.

Moreni, Nicola. 2003. *Pricing American Options: A Variance Reduction Technique for the Longstaff-Schwartz Algorithm*. Technical Report. 2003-256. Champs-sur-Marne: CERMICS-ENPC.

Rasmussen, Nicki S. 2005. Control Variates for Monte Carlo Valuation of American Options. *Journal of Computational Finance* 9: 1. [CrossRef]

Royden, Halsey Lawrence. 1988. *Real Analysis*. Englewood Cliffs: Prentice Hall.

Stentoft, Lars. 2004. Assessing the Least Squares Monte-Carlo Approach to American Option Valuation. *Review of Derivatives Research* 7: 129–68. [CrossRef]

Stentoft, Lars. 2014. Value Function Approximation or Stopping Time Approximation: A Comparison of Two Recent Numerical Methods for American Option Pricing Using Simulation and Regression. *Journal of Computational Finance* 18: 1. [CrossRef]

Su, Yi, and Michael C. Fu. 2000. Importance Sampling in Derivative Securities Pricing. Paper presented at the 2000 Winter Simulation Conference, Orlando, FL, USA, December 10–13. Proceedings (Cat. No. 00CH37165). vol. 1, pp. 587–96.

Tian, Tianhai, and Kevin Burrage. 2002. Accuracy Issues of Monte-Carlo Methods for Valuing American Options. *The ANZIAM Journal* 44: 739–58. [CrossRef]

Vázquez-Abad, Felisa J., and Daniel Dufresne. 1998. Accelerated Simulation for Pricing Asian Options. Paper presented at the 1998 Winter Simulation Conference, Washington, DC, USA, December 13–16. Proceedings (Cat. No. 98CH36274). vol. 2, pp. 1493–500.