Valuing American Options by Simulation: A Simple Least-Squares Approach

Francis A. Longstaff
UCLA

Eduardo S. Schwartz
UCLA

This article presents a simple yet powerful new approach for approximating the value of American options by simulation. The key to this approach is the use of least squares to estimate the conditional expected payoff to the optionholder from continuation. This makes this approach readily applicable in path-dependent and multifactor situations where traditional finite difference techniques cannot be used. We illustrate this technique with several realistic examples including valuing an option when the underlying asset follows a jump-diffusion process and valuing an American swaption in a 20-factor string model of the term structure.

One of the most important problems in option pricing theory is the valuation and optimal exercise of derivatives with American-style exercise features. These types of derivatives are found in all major financial markets including the equity, commodity, foreign exchange, insurance, energy, sovereign, agency, municipal, mortgage, credit, real estate, convertible, swap, and emerging markets. Despite recent advances, however, the valuation and optimal exercise of American options remains one of the most challenging problems in derivatives finance, particularly when more than one factor affects the value of the option. This is primarily because finite difference and binomial techniques become impractical in situations where there are multiple factors.[1]

[1] For example, this explains why virtually all Wall Street firms value and exercise American swaptions using a simple single-factor model despite clear evidence that the term structure is driven by multiple factors.

In this article, we present a simple, yet powerful new approach to approximating the value of American options by simulation. By its nature, simulation is a promising alternative to traditional finite difference and binomial techniques and has many advantages as a framework for valuing, risk managing, and optimally exercising American options. For example, simulation is readily applied when the value of the option depends on multiple factors. Simulation can also be used to value derivatives with both path-dependent and American-exercise features. Simulation allows state variables to follow general stochastic processes such as jump diffusions, as in Merton (1976) and Cox and Ross (1976), non-Markovian processes, as in Heath, Jarrow, and Morton (1992), and even general semimartingales, as in Harrison and Pliska (1981).[2] From a practical perspective, simulation is well suited to parallel computing, which allows significant gains in computational speed and efficiency. Finally, simulation techniques are simple, transparent, and flexible.

To understand the intuition behind this approach, recall that at any exercise time, the holder of an American option optimally compares the payoff from immediate exercise with the expected payoff from continuation, and then exercises if the immediate payoff is higher. Thus the optimal exercise strategy is fundamentally determined by the conditional expectation of the payoff from continuing to keep the option alive. The key insight underlying our approach is that this conditional expectation can be estimated from the cross-sectional information in the simulation by using least squares. Specifically, we regress the ex post realized payoffs from continuation on functions of the values of the state variables. The fitted value from this regression provides a direct estimate of the conditional expectation function. By estimating the conditional expectation function for each exercise date, we obtain a complete specification of the optimal exercise strategy along each path. With this specification, American options can then be valued accurately by simulation. We refer to this technique as the least squares Monte Carlo (LSM) approach.

This approach is easy to implement since nothing more than simple least squares is required. To illustrate this, we present a series of increasingly complex but realistic examples. In the first, we value an American put option in a single-factor setting. In the second, we value an exotic American–Bermuda–Asian option. This option is path dependent and has multifactor features. In the third, we value a cancelable index amortizing swap where the term structure is driven by several factors. This standard fixed-income derivative product has almost pathological path-dependent properties. In each case, the simulation algorithm gives values that are indistinguishable from those given by more computationally intensive finite difference techniques. In the fourth example, we value American options on an asset which follows a jump-diffusion process. This option cannot be valued using standard finite difference techniques. To illustrate the full generality of this approach, the fifth

---

[2] Semimartingales are essentially the broadest class of processes for which stochastic integrals can be defined and standard option pricing theory applied.

example values a deferred American swaption in a 20-factor string model where each point on the interest-rate curve is a separate factor. We also show how the algorithm can be used in a risk-management context by computing the sensitivity of swaption values to each point along the curve.

A number of other recent articles also address the pricing of American options by simulation. In an important early contribution to this literature, Bossaerts (1989) solves for the exercise strategy that maximizes the simulated value of the option. Other important examples of this literature include Tilley (1993), Barraquand and Martineau (1995), Averbukh (1997), Broadie and Glasserman (1997a,b,c), Broadie, Glasserman, and Jain (1997), Raymar and Zwecher (1997), Broadie et al. (1998), Carr (1998), Ibanez and Zapatero (1998), and Garcia (1999). These articles use various stratification or parameterization techniques to approximate the transitional density function or the early exercise boundary. This article takes a fundamentally different approach by focusing directly on the conditional expectation function.

Several recent articles that use an approach similar to ours include Carriere (1996) and Tsitsiklis and Van Roy (1999). Our work, however, differs in a number of ways. For example, neither of these articles take the approach to the level of practical implementation we do in this article. Furthermore, we include in the regression only paths for which the option is in the money. This significantly increases the efficiency of the algorithm and decreases the computational time. In addition, we demonstrate the application of the methodology to complex derivatives with many underlying factors and evaluate the accuracy of the algorithm by comparing our solutions to finite difference approximations.[3]

The remainder of this article is organized as follows. Section 1 presents a simple numerical example of the simulation approach. Section 2 describes the underlying theoretical framework. Sections 3–7 provide specific examples of the application of this approach. Section 8 discusses a number of numerical and implementation issues. Section 9 summarizes the results and contains concluding remarks.

1. A Numerical Example

At the final exercise date, the optimal exercise strategy for an American option is to exercise the option if it is in the money. Prior to the final date, however, the optimal strategy is to compare the immediate exercise value with the expected cash flows from continuing, and then exercise if immediate exercise is more valuable. Thus, the key to optimally exercising an American option is identifying the conditional expected value of continuation. In this approach, we use the cross-sectional information in the simulated paths to

---

[3] Another related article is Keane and Wolpin (1994), which uses regression in a simulation context to solve discrete choice dynamic programming problems.

115

identify the conditional expectation function. This is done by regressing the subsequent realized cash flows from continuation on a set of basis functions of the values of the relevant state variables. The fitted value of this regression is an efficient unbiased estimate of the conditional expectation function and allows us to accurately estimate the optimal stopping rule for the option.

Perhaps the best way to convey the intuition of the LSM approach is through a simple numerical example. Consider an American put option on a share of non-dividend-paying stock. The put option is exercisable at a strike price of 1.10 at times 1, 2, and 3, where time three is the final expiration date of the option. The riskless rate is 6%. For simplicity, we illustrate the algorithm using only eight sample paths for the price of the stock. These sample paths are generated under the risk-neutral measure and are shown in the following matrix.

Stock price paths

Path	$t = 0$	$t = 1$	$t = 2$	$t = 3$
1	1.00	1.09	1.08	1.34
2	1.00	1.16	1.26	1.54
3	1.00	1.22	1.07	1.03
4	1.00	.93	.97	.92
5	1.00	1.11	1.56	1.52
6	1.00	.76	.77	.90
7	1.00	.92	.84	1.01
8	1.00	.88	1.22	1.34

Our objective is to solve for the stopping rule that maximizes the value of the option at each point along each path. Since the algorithm is recursive, however, we first need to compute a number of intermediate matrices. Conditional on not exercising the option before the final expiration date at time 3, the cash flows realized by the optionholder from following the optimal strategy at time 3 are given below.

Cash-flow matrix at time 3

Path	$t = 1$	$t = 2$	$t = 3$
1	—	—	.00
2	—	—	.00
3	—	—	.07
4	—	—	.18
5	—	—	.00
6	—	—	.20
7	—	—	.09
8	—	—	.00

These cash flows are identical to the cash flows that would be received if the option were European rather than American.

116

If the put is in the money at time 2, the optionholder must then decide whether to exercise the option immediately or continue the option's life until the final expiration date at time 3. From the stock-price matrix, there are only five paths for which the option is in the money at time 2. Let X denote the stock prices at time 2 for these five paths and Y denote the corresponding discounted cash flows received at time 3 if the put is not exercised at time 2. We use only in-the-money paths since it allows us to better estimate the conditional expectation function in the region where exercise is relevant and significantly improves the efficiency of the algorithm. The vectors X and Y are given by the nondashed entries below.

<center>Regression at time 2</center>

Path	Y	X
1	$.00 \times .94176$	1.08
2	—	—
3	$.07 \times .94176$	1.07
4	$.18 \times .94176$	.97
5	—	—
6	$.20 \times .94176$	.77
7	$.09 \times .94176$	.84
8	—	—

To estimate the expected cash flow from continuing the option's life conditional on the stock price at time 2, we regress Y on a constant, X, and X^2. This specification is one of the simplest possible; more general specifications are considered later in the article. The resulting conditional expectation function is $E[\,Y\mid X\,] = -1.070 + 2.983X - 1.813X^2$.

With this conditional expectation function, we now compare the value of immediate exercise at time 2, given in the first column below, with the value from continuation, given in the second column below.

<center>Optimal early exercise decision at time 2</center>

Path	Exercise	Continuation
1	.02	.0369
2	—	—
3	.03	.0461
4	.13	.1176
5	—	—
6	.33	.1520
7	.26	.1565
8	—	—

The value of immediate exercise equals the intrinsic value $1.10 - X$ for the in-the-money paths, while the value from continuation is given by substituting X into the conditional expectation function. This comparison implies that

it is optimal to exercise the option at time 2 for the fourth, sixth, and seventh paths. This leads to the following matrix, which shows the cash flows received by the optionholder conditional on not exercising prior to time 2.

Cash-flow matrix at time 2

Path	$t = 1$	$t = 2$	$t = 3$
1	—	.00	.00
2	—	.00	.00
3	—	.00	.07
4	—	.13	.00
5	—	.00	.00
6	—	.33	.00
7	—	.26	.00
8	—	.00	.00

Observe that when the option is exercised at time 2, the cash flow in the final column becomes zero. This is because once the option is exercised there are no further cash flows since the option can only be exercised once.

Proceeding recursively, we next examine whether the option should be exercised at time 1. From the stock price matrix, there are again five paths where the option is in the money at time 1. For these paths, we again define Y as the discounted value of subsequent option cash flows. Note that in defining Y, we use actual realized cash flows along each path; we do not use the conditional expected value of Y estimated at time 2 in defining Y at time 1. As is discussed later, discounting back the conditional expected value rather than actual cash flows can lead to an upward bias in the value of the option.

Since the option can only be exercised once, future cash flows occur at either time 2 or time 3, but not both. Cash flows received at time 2 are discounted back one period to time 1, and any cash flows received at time 3 are discounted back two periods to time 1. Similarly X represents the stock prices at time 1 for the paths where the option is in the money. The vectors X and Y are given by the nondashed elements in the following matrix.

Regression at time 1

Path	Y	X
1	$.00 \times .94176$	1.09
2	—	—
3	—	—
4	$.13 \times .94176$	.93
5	—	—
6	$.33 \times .94176$	.76
7	$.26 \times .94176$	.92
8	$.00 \times .94176$	.88

The conditional expectation function at time 1 is estimated by again regressing Y on a constant, X and X^2. The estimated conditional expectation function is $E[\,Y\mid X\,]=2.038-3.335X+1.356X^2$. Substituting the values of X into this regression gives the estimated conditional expectation function. These estimated continuation values and immediate exercise values at time 1 are given in the first and second columns below. Comparing the two columns shows that exercise at time 1 is optimal for the fourth, sixth, seventh, and eighth paths.

	Optimal early exercise decision at time 1	
Path	Exercise	Continuation
1	.01	.0139
2	—	—
3	—	—
4	.17	.1092
5	—	—
6	.34	.2866
7	.18	.1175
8	.22	.1533

Having identified the exercise strategy at times 1, 2, and 3, the stopping rule can now be represented by the following matrix, where the ones denote exercise dates at which the option is exercised.

	Stopping rule		
Path	$t=1$	$t=2$	$t=3$
1	0	0	0
2	0	0	0
3	0	0	1
4	1	0	0
5	0	0	0
6	1	0	0
7	1	0	0
8	1	0	0

With this specification of the stopping rule, it is now straightforward to determine the cash flows realized by following this stopping rule. This is done by simply exercising the option at the exercise dates where there is a one in the above matrix. This leads to the following option cash

flow matrix.

	Option cash flow matrix		
Path	$t = 1$	$t = 2$	$t = 3$
1	.00	.00	.00
2	.00	.00	.00
3	.00	.00	.07
4	.17	.00	.00
5	.00	.00	.00
6	.34	.00	.00
7	.18	.00	.00
8	.22	.00	.00

Having identified the cash flows generated by the American put at each date along each path, the option can now be valued by discounting each cash flow in the option cash flow matrix back to time zero, and averaging over all paths. Applying this procedure results in a value of .1144 for the American put. This is roughly twice the value of .0564 for the European put obtained by discounting back the cash flows at time 3 from the first cash flow matrix.

Although very stylized, this example illustrates how least squares can use the cross-sectional information in the simulated paths to estimate the conditional expectation function. In turn, the conditional expectation function is used to identify the exercise decision that maximizes the value of the option at each date along each path. As shown by this example, the LSM approach is easily implemented since nothing more than simple regression is involved.

2. The Valuation Algorithm

In this section we describe the general LSM algorithm. The valuation framework underlying the LSM algorithm is based on the general derivative pricing paradigm of Black and Scholes (1973), Merton (1973), Harrison and Kreps (1979), Harrison and Pliska (1981), Cox, Ingersoll, and Ross (1985), Heath, Jarrow, and Morton (1992), and others. We also present several convergence results for the algorithm.

2.1 The valuation framework

We assume an underlying complete probability space $(\Omega, \mathcal{F}, P)$ and finite time horizon $[0, T]$, where the state space Ω is the set of all possible realizations of the stochastic economy between time 0 and T and has typical element ω representing a sample path, $\mathcal{F}$ is the sigma field of distinguishable events at time T, and P is a probability measure defined on the elements of $\mathcal{F}$. We define $F = \{\mathcal{F}_t; t \in [0, T]\}$ to be the augmented filtration generated by the relevant price processes for the securities in the economy, and assume that $\mathcal{F}_T = \mathcal{F}$. Consistent with the no-arbitrage paradigm, we assume the existence of an equivalent martingale measure Q for this economy.

We are interested in valuing American-style derivative securities with random cash flows which may occur during $[0, T]$. We restrict our attention to

derivatives with payoffs that are elements of the space of square-integrable or finite-variance functions $L^2(\Omega, \mathcal{F}, Q)$. Standard results such as Bensoussan (1984) and Karatzas (1988) imply that the value of an American option can be represented by the Snell envelope; the value of an American option equals the maximized value of the discounted cash flows from the option, where the maximum is taken over all stopping times with respect to the filtration F. We introduce the notation $C(\omega, s; t, T)$ to denote the path of cash flows generated by the option, conditional on the option not being exercised at or prior to time t and on the optionholder following the optimal stopping strategy for all s, $t < s \leq T$. This function is analogous to the intermediate cash-flow matrices used in the previous section.

The objective of the LSM algorithm is to provide a pathwise approximation to the optimal stopping rule that maximizes the value of the American option. To convey the intuition behind the LSM algorithm, we focus the discussion on the case where the American option can only be exercised at the K discrete times $0 < t_1 \leq t_2 \leq t_3 \leq \cdots \leq t_K = T$, and consider the optimal stopping policy at each exercise date. In practice, of course, many American options are continuously exercisable; the LSM algorithm can be used to approximate the value of these options by taking K to be sufficiently large.

At the final expiration date of the option, the investor exercises the option if it is in the money, or allows it to expire if it is out of the money. At exercise time t_k prior to the final expiration date, however, the optionholder must choose whether to exercise immediately or to continue the life of the option and revisit the exercise decision at the next exercise time. The value of the option is maximized pathwise, and hence unconditionally, if the investor exercises as soon as the immediate exercise value is greater than or equal to the value of continuation.[4]

At time t_k, the cash flow from immediate exercise is known to the investor, and the value of immediate exercise simply equals this cash flow. The cash flows from continuation, of course, are not known at time t_k. No-arbitrage valuation theory, however, implies that the value of continuation, or equivalently, the value of the option assuming that it cannot be exercised until after t_k, is given by taking the expectation of the remaining discounted cash flows $C(\omega, s; t_k, T)$ with respect to the risk-neutral pricing measure Q. Specifically, at time t_k, the value of continuation $F(\omega; t_k)$ can be expressed as

$$F(\omega; t_k) = E_Q\left[\sum_{j=k+1}^{K} \exp\left(-\int_{t_k}^{t_j} r(\omega, s)ds\right) C(\omega, t_j; t_k, T) \mid \mathcal{F}_{t_k} \right], \qquad (1)$$

---

[4] For a discussion of optimal exercise policies for American options, see Duffie (1996) or Lamberton and Lapeyre (1996). Bossaerts (1989) directly uses this maximization property in developing simulation estimates of American option prices by parameterizing the stopping rule and then solving for the parameters that maximize the value of the option.

where $r(\omega, t)$ is the (possibly stochastic) riskless discount rate, and the expectation is taken conditional on the information set $\mathcal{F}_{t_k}$ at time t_k. With this representation, the problem of optimal exercise reduces to comparing the immediate exercise value with this conditional expectation, and then exercising as soon as the immediate exercise value is positive and greater than or equal to the conditional expectation.

2.2 The LSM algorithm

The LSM approach uses least squares to approximate the conditional expectation function at $t_{K-1}, t_{K-2}, \ldots, t_1$. We work backwards since the path of cash flows $C(\omega, s; t, T)$ generated by the option is defined recursively; $C(\omega, s; t_k, T)$ can differ from $C(\omega, s; t_{k+1}, T)$ since it may be optimal to stop at time t_{k+1}, thereby changing all subsequent cash flows along a realized path ω. Specifically, at time t_{K-1}, we assume that the unknown functional form of $F(\omega; t_{K-1})$ in Equation (1) can be represented as a linear combination of a countable set of $\mathcal{F}_{t_{K-1}}$-measurable basis functions.

This assumption can be formally justified, for example, when the conditional expectation is an element of the L^2 space of square-integrable functions relative to some measure. Since L^2 is a Hilbert space, it has a countable orthonormal basis and the conditional expectation can be represented as a linear function of the elements of the basis.[5] As an example, assume that X is the value of the asset underlying the option and that X follows a Markov process.[6] One possible choice of basis functions is the set of (weighted) Laguerre polynomials:

$$L_0(X) = \exp(-X/2), \tag{2}$$

$$L_1(X) = \exp(-X/2)\,(1 - X), \tag{3}$$

$$L_2(X) = \exp(-X/2)\,(1 - 2X + X^2/2), \tag{4}$$

$$L_n(X) = \exp(-X/2)\frac{e^X}{n!}\frac{d^n}{dX^n}(X^n e^{-X}). \tag{5}$$

With this specification, $F(\omega; t_{K-1})$ can be represented as

$$F(\omega; t_{K-1}) = \sum_{j=0}^{\infty} a_j\, L_j(X), \tag{6}$$

where the a_j coefficients are constants. Other types of basis functions include the Hermite, Legendre, Chebyshev, Gegenbauer, and Jacobi polynomials.[7]

---

[5] For a discussion of Hilbert space theory and Hilbert space representations of square-integrable functions, see Royden (1968).

[6] For Markovian problems, only current values of the state variables are necessary. For non-Markovian problems, both current and past realizations of the state variables can be included in the basis functions and the regressions.

[7] These functions are described in Chapter 22 of Abramowitz and Stegun (1970).

Numerical tests indicate that Fourier or trigonometric series and even simple powers of the state variables also give accurate results.

To implement the LSM approach, we approximate $F(\omega; t_{K-1})$ using the first $M < \infty$ basis functions, and denote this approximation $F_M(\omega; t_{K-1})$. Once this subset of basis functions has been specified, $F_M(\omega; t_{K-1})$ is estimated by projecting or regressing the discounted values of $C(\omega, s; t_{K-1}, T)$ onto the basis functions for the paths where the option is in the money at time t_{K-1}. We use only in-the-money paths in the estimation since the exercise decision is only relevant when the option is in the money. By focusing on the in-the-money paths, we limit the region over which the conditional expectation must be estimated, and far fewer basis functions are needed to obtain an accurate approximation to the conditional expectation function.[8] Since the values of the basis functions are independently and identically distributed across paths, weak assumptions about the existence of moments allow us to use Theorem 3.5 of White (1984) to show that the fitted value of this regression $\widehat{F}_M(\omega; t_{K-1})$ converges in mean square and in probability to $F_M(\omega; t_{K-1})$ as the number N of (in-the-money) paths in the simulation goes to infinity. Furthermore, Theorem 1.2.1 of Amemiya (1985) implies that $\widehat{F}_M(\omega; t_{K-1})$ is the best linear unbiased estimator of $F_M(\omega; t_{K-1})$ based on a mean-squared metric.

Once the conditional expectation function at time t_{K-1} is estimated, we can determine whether early exercise at time t_{K-1} is optimal for an in-the-money path ω by comparing the immediate exercise value with $\widehat{F}_M(\omega; t_{K-1})$, and repeating for each in-the-money path. Once the exercise decision is identified, the option cash flow paths $C(\omega, s; t_{K-2}, T)$ can then be approximated. The recursion proceeds by rolling back to time t_{K-2} and repeating the procedure until the exercise decisions at each exercise time along each path have been determined. The American option is then valued by starting at time zero, moving forward along each path until the first stopping time occurs, discounting the resulting cash flow from exercise back to time zero, and then taking the average over all paths ω.

When there are two state variables X and Y, the set of basis functions should include terms in X and in Y, as well as cross-products of these terms. Similarly for higher-dimensional problems. Intuitively this seems to suggest that the number of basis functions needed grows exponentially with the dimensionality of the problem. In actuality, however, there may be reasons why the number of basis functions necessary to obtain a desired level of convergence might grow at a slower rate. As an example, Judd (1998)

---

[8] We conducted a variety of numerical experiments which indicated that if all paths are used, more than two or three times as many basis functions may be needed to obtain the same level of accuracy as obtained by the estimator based on in-the-money paths. Intuitively this makes sense since we are interested in estimating the expectation conditional on the current state and the event that the option is in the money. By using all paths, and hence, not conditioning on the event that the option is in the money, we obtain estimates of the conditional expectation function which have larger standard errors than those obtained by using all of the conditioning information.

shows that by using sets of complete polynomials, kth degree convergence is obtained asymptotically using a number of terms that grows only polynomially with the dimension of the problem. Similar results are also well known in the neural-network literature; as examples, see Barron (1993) and Refenes, Burgess, and Bentz (1997). In the numerical results given later in the article, we find that the number of basis functions needed to obtain convergence appears to grow much more slowly than exponentially. Our experience suggests that the number of basis functions necessary to approximate the conditional expectation function may be very manageable even for high-dimensional problems.

2.3 Convergence results

The LSM algorithm provides a simple and elegant way of approximating the optimal early exercise strategy for an American-style option. While the ultimate test of the algorithm is how well it performs using a realistic number of paths and basis functions, it is also useful to examine what can be said about the theoretical convergence of the algorithm to the true value $V(X)$ of the American option.

The first convergence result addresses the bias of the LSM algorithm and is applicable even when the American option is continuously exercisable.

Proposition 1. *For any finite choice of M, K, and vector $\theta \in R^{M \times (K-1)}$ representing the coefficients for the M basis functions at each of the $K-1$ early exercise dates, let $LSM(\omega; M, K)$ denote the discounted cash flow resulting from following the LSM rule of exercising when the immediate exercise value is positive and greater than or equal to $\widehat{F}_M(\omega_i; t_k)$ as defined by θ. Then the following inequality holds almost surely,*

$$V(X) \geq \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} LSM(\omega_i; M, K).$$

Proof. See the appendix. ∎

The intuition for this result is easily understood. The LSM algorithm results in a stopping rule for an American-style option. The value of an American-style option, however, is based on the stopping rule that maximizes the value of the option; all other stopping rules, including the stopping rule implied by the LSM algorithm, result in values less than or equal to that implied by the optimal stopping rule.

This result is particularly useful since it provides an objective criterion for convergence. For example, this criterion provides guidance in determining the number of basis functions needed to obtain an accurate approximation; simply increase M until the value implied by the LSM algorithm no longer

increases. This useful and important property is not shared by algorithms that simply discount back functions based on the estimated continuation value.[9]

By its nature, providing a general convergence result for the LSM algorithm is difficult since we need to consider limits as the number of discretization points K, the number of basis functions M, and the number of paths N go to infinity. In addition, we need to consider the effects of propagating the estimating stopping rule backwards through time from t_{K-1} to t_1. In the case where the American option can only be exercised at $K = 2$ discrete points in time, however, convergence of the algorithm is more easily demonstrated. As an example, consider the following proposition.

Proposition 2. *Assume that the value of an American option depends on a single state variable X with support on $(0, \infty)$ which follows a Markov process. Assume further that the option can only be exercised at times t_1 and t_2, and that the conditional expectation function $F(\omega; t_1)$ is absolutely continuous and*

$$\int_0^\infty e^{-X} F^2(\omega; t_1) dX \ < \ \infty,$$

$$\int_0^\infty e^{-X} F_X^2(\omega; t_1) dX \ < \infty.$$

Then for any $\epsilon > 0$, there exists an $M < \infty$ such that

$$\lim_{N \to \infty} \Pr\left[|\ V(X) - \frac{1}{N} \sum_{i=1}^N LSM(\omega_i; M, K)\ | > \epsilon \right] = 0.$$

Proof. See the appendix. ∎

Intuitively this result means that by selecting M large enough and letting $N \to \infty$, the LSM algorithm results in a value for the American option within ϵ of the true value. Thus the LSM algorithm converges to any desired degree of accuracy since ϵ is arbitrary. The key to this result is that the convergence of $F_M(\omega; t_1)$ to $F(\omega; t_1)$ is uniform on $(0, \infty)$ when the indicated integrability conditions are met. This bounds the maximum error in estimating the conditional expectation, which in turn, bounds the maximum pricing error. An important implication of this result is that the number of basis functions needed to obtain a desired level of accuracy need not go to infinity as $N \to \infty$. While this proposition is limited to one-dimensional settings, we conjecture that similar results can be obtained for higher-dimensional problems by finding conditions under which uniform convergence occurs.

---

[9] For example, if the American option were valued by taking the maximum of the immediate exercise value and the estimated continuation value, and discounting this value back, the resulting American option value could be severely upward biased. This bias arises since the maximum operator is convex; measurement error in the estimated continuation value results in the maximum operator being upward biased. We are indebted to Peter Bossaerts for making this point.

3. Valuing American Put Options

Earlier we used a stylized example to illustrate how this approach could be applied to the valuation of American put options. In this section we present an in-depth example of the application of the LSM algorithm to American put options

Assume that we are interested in pricing an American-style put option on a share of stock, where the risk-neutral stock price process follows the stochastic differential equation

$$dS = rSdt + \sigma s dZ, \tag{7}$$

and where r and σ are constants, Z is a standard Brownian motion, and the stock does not pay dividends. Furthermore, assume that the option is excercisable 50 times per year at a strike price of K up to and including the final expiration date T of the option. This type of discrete American-style exercise feature is also sometimes termed a Bermuda exercise feature. As the set of basis functions, we use a constant and the first three Laguerre polynomials as given in Equations (2)–(4). Thus we regress discounted realized cash flows on a constant and three nonlinear functions of the stock price. Since we use linear regression to estimate the conditional expectation function, it is straightforward to add additional basis functions as explanatory variables in the regression if needed. Using more than three basis functions, however, does not change the numerical results; three basis functions are sufficient to obtain effective convergence of the algorithm in this example.

To illustrate the results, Table 1 reports the values of the early exercise option implied by both the finite difference and LSM techniques. The value of the early exercise option is the difference between the American and European put values. The European put value is given by the Black–Scholes formula. In this article we focus primarily on the early exercise value since it is the most difficult component of an American options's value to determine; the European component of an American option's value is much easier to identify.

The finite difference results reported in Table 1 are obtained from an implicit finite difference scheme with 40,000 time steps per year and 1,000 steps for the stock price. The partial differential equation satisfied by the put price $P(S, t)$ is

$$(\sigma^2 S^2/2)P_{SS} + rSP_S - rP + P_T = 0, \tag{8}$$

subject to the expiration condition $P(S, T) = \max(0, K - S_T)$. The implicit finite difference results were also compared with the results from an explicit finite difference algorithm; the two finite difference techniques resulted in values that were generally within one cent of each other. The LSM estimates

Table 1

S	σ	T	Finite difference American	Closed form European	Early exercise Value	Simulated American	(s.e.)	Closed form European	Early exercise value	Difference in early exercise value
36	.20	1	4.478	3.844	.634	4.472	(.010)	3.844	.628	.006
36	.20	2	4.840	3.763	1.077	4.821	(.012)	3.763	1.058	.019
36	.40	1	7.101	6.711	.390	7.091	(.020)	6.711	.380	.010
36	.40	2	8.508	7.700	.808	8.488	(.024)	7.700	.788	.020
38	.20	1	3.250	2.852	.398	3.244	(.009)	2.852	.392	.006
38	.20	2	3.745	2.991	.754	3.735	(.011)	2.991	.744	.010
38	.40	1	6.148	5.834	.314	6.139	(.019)	5.834	.305	.009
38	.40	2	7.670	6.979	.691	7.669	(.022)	6.979	.690	.001
40	.20	1	2.314	2.066	.248	2.313	(.009)	2.066	.247	.001
40	.20	2	2.885	2.356	.529	2.879	(.010)	2.356	.523	.006
40	.40	1	5.312	5.060	.252	5.308	(.018)	5.060	.248	.004
40	.40	2	6.920	6.326	.594	6.921	(.022)	6.326	.595	−.001
42	.20	1	1.617	1.465	.152	1.617	(.007)	1.465	.152	.000
42	.20	2	2.212	1.841	.371	2.206	(.010)	1.841	.365	.006
42	.40	1	4.582	4.379	.203	4.588	(.017)	4.379	.209	−.006
42	.40	2	6.248	5.736	.512	6.243	(.021)	5.736	.507	.005
44	.20	1	1.110	1.017	.093	1.118	(.007)	1.017	.101	−.008
44	.20	2	1.690	1.429	.261	1.675	(.009)	1.429	.246	.015
44	.40	1	3.948	3.783	.165	3.957	(.017)	3.783	.174	−.009
44	.40	2	5.647	5.202	.445	5.622	(.021)	5.202	.420	.025

Comparison of the finite difference and simulation values for the early exercise option in an American-style put option on a share of stock, where the option is exercisable 50 times per year. The early exercise value is the difference between the American and European put values. In this comparison, the strike price of the put is 40, the short-term interest rate is .06, and the underlying stock price S, the volatility of returns σ, and the number of years until the final expiration of the option T are as indicated. The European option values are based on the closed-form Black–Scholes formula. The simulation is based on 100,000 (50,000 plus 50,000 antithetic) paths for the stock-price process. The standard errors of the simulation estimates (s.e.) are given in parentheses.

are based on 100,000 (50,000 plus 50,000 antithetic) paths using 50 exercise points per year.

As shown, the differences between the finite difference and LSM algorithms are typically very small. Of the 20 differences shown in Table 1, 16 are less than or equal to one cent in absolute value. The standard errors for the simulated values range from 0.7 to 2.4 cents, which is well within the market bid-ask spread for these types of options.[10] In addition, the differences are both positive and negative. These results suggest that the LSM algorithm is able to approximate closely the finite-difference values.

Broadie ad Glasserman (1997a,b), Raymar and Zwecher (1997), Garcia (1999), and others suggest an interesting diagnostic test for the convergence of a simulation algorithm. Essentially the stopping rule is estimated from one set of paths and then applied to another set of paths. In our context, this can be implemented by estimating the conditional expectation regressions

---

[10] Exchange traded stock option premiums are typically quoted in sixteenths or eighths of a dollar. The bid-ask spread for an at-the-money option would generally be some multiple of these fractions.

from one set of paths and then applying the regression functions to an out-of-sample set of paths. A successful algorithm should lead to out-of-sample values that closely approximate the in-sample values for the option.[11]

The results from these diagnostic tests are shown in Table 2. For selected sets of parameters from Table 1, we estimate the regressions in sample, value the option using the in-sample LSM procedure, and then value the option out of sample using the in-sample regression parameters but different paths. We repeat this process for five different initial seeds of the random number generator; the five rows for each example shown in Table 2 correspond to different initial seeds. As shown, the in-sample and out-of-sample values are virtually identical. The differences between the in-sample and out-of-sample values are virtually identical. The differences between the in-sample and out-of-sample values are both positive and negative and only 5% of the values are larger than two standard errors. This provides strong support for the accuracy of the algorithm. Given these results, we recommend using the LSM algorithm in sample in order to minimize computational time.

4. Valuing an American–Bermuda–Asian Option

In this section we apply the LSM algorithm to a more exotic path-dependent option. In particular, we consider a call option on the average price of a stock over some horizon, where the call option can be exercised at any time after some initial lockout period. Thus this option is an Asian option since it is an option on an average, and has both a Bermuda and American exercise feature; an American–Bermuda–Asian option.

Define the current valuation date as time 0. We assume that the option has a final expiration date of $T = 2$, and that the option can be exercised at any time after $t = .25$ by payment of the strike price K. The underlying average A_t, $.25 \leq t \leq T$, is the continuous arithmetic average of the underlying stock price during the period three months prior to the valuation date (a three-month lookback) to time t. Thus the cash flow from exercising the option at time t is $\max(0, A_t - K)$. The risk-neutral dynamics for the stock price are the same as in the previous section.

This option is particularly complex because it not only has an American exercise feature, but the cash flow from exercise is path dependent since A_t depends on the path of the stock price over the averaging window. In general, these types of problems are very difficult to solve using finite difference techniques. In this case, we can value the option by finite difference techniques by transforming the problem from a path-dependent one to a Markovian problem. This is done by introducing the average to date as a second state variable

---

[11] We are grateful to the referee for suggesting this diagnostic test and for providing some results about the performance of the LSM algorithm.

Table 2

S	σ	T	LSM in sample		LSM out of sample		Difference
			Value	(s.e)	Value	(s.e.)	
36	.20	1	4.472	(.010)	4.476	(.010)	−.004
			4.463	(.010)	4.474	(.010)	−.011
			4.467	(.010)	4.480	(.010)	−.013
			4.480	(.010)	4.476	(.010)	.004
			4.468	(.010)	4.469	(.010)	−.001
36	.40	1	7.091	(.020)	7.102	(.020)	−.011
			7.095	(.020)	7.094	(.020)	.001
			7.087	(.020)	7.087	(.020)	.000
			7.097	(.020)	7.095	(.020)	.002
			7.097	(.020)	7.101	(.020)	−.004
36	.20	2	4.821	(.012)	4.818	(.012)	.003
			4.819	(.012)	4.833	(.012)	−.014
			4.820	(.012)	4.829	(.012)	.009
			4.827	(.012)	4.825	(.012)	.002
			4.827	(.012)	4.829	(.012)	−.002
36	.40	2	8.488	(.024)	8.487	(.024)	.001
			8.485	(.024)	8.478	(.024)	.007
			8.483	(.024)	8.483	(.024)	.000
			8.495	(.023)	8.498	(.024)	−.003
			8.493	(.024)	8.491	(.024)	.002
Mean							−.002
44	.20	1	1.118	(.007)	1.102	(.007)	.016
			1.108	(.007)	1.114	(.007)	−.006
			1.115	(.007)	1.099	(.007)	.016
			1.108	(.007)	1.097	(.007)	.011
			1.114	(.007)	1.107	(.007)	.007
44	.40	1	3.957	(.017)	3.927	(.017)	.030
			3.941	(.017)	3.976	(.017)	−.035
			3.953	(.017)	3.924	(.017)	.029
			3.939	(.017)	3.924	(.017)	.015
			3.953	(.017)	3.945	(.017)	.008
44	.20	2	1.675	(.009)	1.691	(.009)	−.016
			1.673	(.009)	1.682	(.009)	−.009
			1.687	(.009)	1.702	(.009)	−.015
			1.671	(.009)	1.674	(.009)	−.003
			1.696	(.009)	1.688	(.009)	.008
44	.40	2	5.622	(.021)	5.644	(.021)	−.022
			5.637	(.021)	5.637	(.021)	.000
			5.628	(.021)	5.652	(.021)	−.024
			5.615	(.021)	5.632	(.021)	−.017
			5.649	(.021)	5.628	(.021)	.021
Mean							.001

Comparison of the in-sample and out-of-sample LSM estimates of the value of an American-style put option on a share of stock, where the option is exercisable 50 times per year. In this comparison, the strike price of the put is 40, and the short-term interest rate is .06. The underlying stock price S, the volatility of returns σ, and the number of years until the final expiration date of the option T are as indicated. The LSM valuations for each of the indicated options are repeated five times with different initial seeds for the random number generator; the five rows for each of the options are based on a different seed value. The in-sample and out-of-sample comparisons are each based on 100,000 (50,000 plus 50,000 antithetic) paths of the stock-price process. The standard errors of the simulation estimates (s.e.) are given in parentheses.

Table 3

		Finite difference			Simulation					
				Early exercise value					Early exercise value	Difference in early exercise value
A	S	American	European		American	(s.e)	European	(s.e)		
90	80	.949	.949	.000	.961	(.016)	.951	(.016)	.010	−.010
90	90	3.267	3.230	.037	3.309	(.030)	3.233	(.030)	.076	−.039
90	100	7.889	7.569	.320	7.886	(.046)	7.573	(.046)	.313	.007
90	110	14.538	13.775	.763	14.518	(.059)	13.783	(.061)	.735	.028
90	120	22.423	21.196	1.227	22.378	(.068)	21.201	(.071)	1.177	.050
100	80	1.108	1.082	.026	1.101	(.017)	1.085	(.017)	.016	.010
100	90	3.710	3.567	.143	3.700	(.032)	3.570	(.032)	.130	.013
100	100	8.658	8.151	.507	8.669	(.047)	8.156	(.047)	.513	−.006
100	110	15.717	14.558	1.159	15.703	(.059)	14.565	(.061)	1.138	.021
100	120	23.811	22.097	1.714	23.775	(.066)	22.101	(.072)	1.674	.040
110	80	1.288	1.232	.056	1.265	(.018)	1.235	(.018)	.030	.026
110	90	4.136	3.933	.203	4.186	(.033)	3.936	(.033)	.250	−.047
110	100	9.821	8.764	1.057	9.830	(.046)	8.768	(.049)	1.062	−.005
110	110	17.399	15.361	2.038	17.362	(.056)	15.368	(.062)	1.994	.044
110	120	25.453	23.009	2.444	25.406	(.064)	23.013	(.073)	2.393	.051

Comparison of the finite difference and simulation values for the early exercise option in an American–Bermuda–Asian option on a share of stock. The early exercise value is the difference in the value of the American–Bermuda–Asian option and its European counterpart. In this example, the strike price is 100, the short-term interest rate is .06, the initial average value of the stock is A, the underlying stock price is S, the volatility of returns is .20, and the final expiration of the option is in two years. The average stock price is computed over the period beginning three months before the valuation date to the exercise date. The option is not exercisable until three months after the valuation date. The simulation is based on 50,000 (25,000 plus 25,000 antithetic) paths for the stock-price process with 100 discretization points per year. The standard errors of the simulation estimates (s.e.) are given in parantheses.

in the problem. Consequently the option price $H(S, A, t)$ is the solution of the following two-dimensional partial differential equation

$$(\sigma^2 S^2/2) H_{SS} + r S H_S + \frac{1}{.25 + t}(S - A) H_A - r H + H_t = 0, \qquad (9)$$

subject to the expiration condition $H(S, A, T) = \max(0, A_T - K)$. Note that the path dependence of the option payoff does not pose any difficulties to the simulation-based LSM algorithm.

Table 3 compares the numerical results from valuing this option by finite difference techniques with those obtained by the LSM approach. In this example, we compute both the value of the American–Bermuda–Asian option and its European counterpart, and focus primarily on the difference which represents the value of the early exercise option. The European counterpart of the American–Bermuda–Asian option is an option on the average stock price which can only be exercised at the final option maturity date $T = 2$. In this example, as well as in later examples in the article, we use the same paths to price the European option that is used to value the American option.

The finite difference results are obtained using an alternating directions implicit (ADI) algorithm with 10,000 time steps per year and 200 steps in both the stock price and the average stock price. The results were checked against a standard explicit finite difference scheme with a similar number of

steps for the stock price and the average stock price. The finite difference algorithms result in values that are typically within three cents per \$100 notional. The LSM results are based on 50,000 (25,000 plus 25,000 anti-thetic) paths and use 100 discretization points per year to approximate the continuous exercise feature of the option. As basis functions in the regressions, we use a constant, the first two Laguerre polynomials evaluated at the stock price, the first two Laguerre polynomials evaluated at the average stock price, and the cross products of these Laguerre polynomials up to third-order terms. Thus we use a total of eight basis functions in the regressions.[12]

As shown in Table 3, the finite difference and LSM results are very similar. The differences in the early exercise values are typically less than two or three cents per \$100 notional value. The differences are again both positive and negative; there is no evidence that the LSM algorithm systematically undervalues the early exercise option. The differences in the early exercise values are also small relative to the level of the early exercise value, and very small relative to the level of the American and European option values. These differences would likely be well within the bid-ask spread or other transaction cost bounds.

5. Valuing Cancelable Index Amortizing Swaps

This section uses the LSM approach to value a cancelable index amortizing swap in a multifactor term structure model. Index amortizing swaps have been widely used on Wall Street in recent years and are among the most difficult types of structured interest-rate derivative products to value and risk manage. The reason for this is that the notional amount of these swaps declines over time in a complex way. For example, index amortizing swaps often have notional amounts that amortize on the basis of a nonlinear function of a constant maturity Treasury (CMT) or constant maturity swap (CMS) rate. This stochastic amortization property makes these derivatives highly path dependent. These swaps become even more complex when one of the counterparties has the right to cancel the swap at any time; a cancelable index amortizing swap consists of both an index amortizing swap and an American-style cancellation option. Index amortizing swaps are widely used to hedge or mimic the cash flows from mortgages; the amortization feature of an index amortizing swap is typically patterned after a mortgage prepayment model.

To simplify the example, we focus on a specific cancelable index amortizing swap with a five-year final maturity. We assume that the counterparty with the right to cancel receives a fixed coupon c and pays the floating Libor rate in the swap. We assume that the fixed coupon is received continuously and the floating rate is paid continuously on an actual/actual basis.

---

[12] Numerical tests indicated that adding additional basis functions had little or no effect on the results; using eight basis functions was sufficient to obtain effective convergence.

The notional balance on which the fixed and floating cash flows are based is initially \$100, but amortizes continuously on the basis of the 10-year par swap rate, CMS 10. Let I_t denote the notional balance of the swap at time t. The dynamics of I_t are given by

$$dI = -f(\text{CMS } 10)dt, \tag{10}$$

where $f(u) = .00, u \geq .07$, $f(.06) = .10$, $f(.05) = .50$, $f(v) = 4.00$, $v \leq .04$. When CMS 10 is between .04 and .07, the function $f(\text{CMS } 10)$ is linearly interpolated between the two closest points in this schedule. For example, if CMS 10 = .0543, $f(\text{CMS } 10) = .328$. Note that $f(\text{CMS } 10)$ is a rate; $f(\text{CMS } 10) = 4.00$ implies that the swap is amortizing at a rate that would completely amortize the swap in three months. The counterparty with the right to cancel can choose to cancel the swap at any time; canceling the swap terminates all future cash flows from the swap to either party. Since the cash flows accrue and pay on a continuous basis, there is no lump sum accrued coupon or interest payment made when the swap is canceled.[13]

In this example, we assume that the swap term structure is determined by a simple two-factor Vasicek (1977) type of model.[14] In particular, we assume that the instantaneous Libor rate r equals the sum of two state variables, $r = X + Y$. The risk-neutral dynamics of X and Y are assumed to be

$$dX = (\alpha - \beta X)dt + \sigma \, dZ_1, \tag{11}$$

$$dY = (\gamma - \eta Y)dt + s \, dZ_2, \tag{12}$$

where α, β, σ, γ, η, and s are constants and Z_1 and Z_2 are independent Brownian motions. In this framework, the value of a zero-coupon bond $D(X, Y, T)$ with maturity T is given by

$$D(X, Y, T) = \exp\big(-M(X, Y, T) + V(T)/2\big), \tag{13}$$

---

[13] In practice, index amortizing swaps typically follow the swap market convention of exchanging payments on a quarterly or semiannual cycle, where the floating payment is determined at the beginning of the cycle and paid at the end of the cycle. Cancelable index amortizing swaps typically can only be canceled on a coupon payment date, and only after exchanging any accrued fixed or floating payments.

[14] In this example, and in the swaption example in Section 8, we make the standard simplifying assumption that the swap curve can be modeled as if it were a risk-free term structure. In reality, Libor rates incorporate some small default-risk component. Since both legs of the swap are discounted using the same curve, however, the net effect on the value of the swap is typically very small.

where

$$M(X, Y, T) = \frac{\alpha}{\beta}T + \left(X - \frac{\alpha}{\beta}\right)\frac{(1 - \exp(-\beta T))}{\beta}$$
$$+ \frac{\gamma}{\eta}T + \left(Y - \frac{\gamma}{\eta}\right)\frac{(1 - \exp(-\eta T))}{\eta}$$

$$V(T) = \frac{\sigma^2}{\beta^2}\left(T - \frac{2}{\beta}(1 - \exp(-\beta T)) + \frac{1}{2\beta}(1 - \exp(-2\beta T))\right)$$
$$+ \frac{s^2}{\eta^2}\left(T - \frac{2}{\eta}(1 - \exp(-\eta T)) + \frac{1}{2\eta}(1 - \exp(-2\eta T))\right)$$

The CMS 10 rate, conditional on the current values of X and Y, is given by

$$\text{CMS } 10 = 2 \times \left(\frac{1 - D(X, Y, 10)}{\sum_{i=1}^{20} D(X, Y, i/2)}\right). \quad (14)$$

In this model, the value of the cancelable index amortizing swap depends on three state variables; the two factors X and Y as well as the current notional amount I. The swap $S(X, Y, I, t)$ can be valued by finite difference techniques by solving the following three-dimensional partial differential equation implied by the dynamics for X, Y, and I.

$$(\sigma^2/2)S_{XX} + (s^2/2)S_{YY} + (\alpha - \beta X)S_X + (\gamma - \eta Y)S_Y$$
$$- f(\text{CMS } 10)S_I - (X + Y)S + (c - X - Y)I + S_t = 0, \quad (15)$$

subject to the conditions $S(X, Y, I, T) = 0$ and $S(X, Y, 0, t) = 0$. Similarly, the swap can also be valued by the LSM algorithm by simulating paths of X and Y and keeping track of the notional balance along each path. The parameter values used in this example are chosen to approximate a current term structure and cap volatility curve.[15]

Table 4 presents the numerical results from the finite difference and LSM valuation of the cancellation option on the underlying index amortizing swap. The value of the cancellation option is the difference between the value of the cancelable index amortizing swap and the underlying noncancelable index amortizing swap. The table reports the results for a range of different values of the fixed coupon paid on the swap. The finite difference methodology is an implementation of a successive overrelaxation technique similar to that described in Press et al. (1992). The finite difference algorithm uses 50 steps for X, 40 steps for Y, and 15 steps for I. The LSM algorithm is based on

---

[15] The parameter values used in the example are $\alpha = .001$, $\beta = .1$, $\gamma = .0525$, $\eta = 1.00$, $\sigma = .006951$, $s = .00867$. The initial values of X and Y are .002 and .050.

Table 4

Coupon	Finite difference			Simulation					Difference in cancellation option
	Cancelable	Noncancelable	Cancellation option	Cancelable	(s.e.)	Noncancelable	(s.e.)	Cancellation option	
.0500	.072	−1.928	2.000	.043	(.003)	−1.955	(.033)	1.998	.002
.0510	.118	−1.682	1.800	.087	(.004)	−1.715	(.032)	1.802	−.002
.0520	.182	−1.435	1.617	.141	(.005)	−1.475	(.031)	1.616	.001
.0530	.254	−1.189	1.443	.206	(.005)	−1.235	(.030)	1.441	.002
.0540	.342	−.942	1.284	.284	(.006)	−.995	(.029)	1.279	.005
.0550	.441	−.696	1.137	.377	(.007)	−.755	(.028)	1.132	.005
.0560	.551	−.449	1.000	.480	(.008)	−.515	(.027)	.995	.005
.0570	.676	−.203	.879	.593	(.009)	−.275	(.026)	.868	.011
.0580	.811	.044	.767	.720	(.009)	−.035	(.025)	.755	.012
.0590	.957	.290	.667	.861	(.010)	.205	(.024)	.656	.011
.0600	1.115	.537	.578	1.011	(.011)	.445	(.023)	.566	.012

Comparison of the finite difference and simulation values for a cancellation option on an index amortizing swap. The value of the cancellation option is the difference between the values of the cancelable and noncancelable index amortizing swaps. Coupon denotes the coupon rate for the fixed leg of the swap. Values are given per $100 notional amount. The simulation is based on 5,000 (2,500 plus 2,500 antithetic) paths with 12 discretization points per year. The standard errors of the simulation estimates (s.e) are given in parentheses.

5,000 (2,500 plus 2,500 antithetic in both X and Y) paths. As basis functions, we use a constant, the first three powers of the value of the underlying noncancelable swap, X, X^2, Y, Y^2, and XY. This results in a total of nine basis functions.[16]

As shown in Table 4, the two valuation approaches produce very similar numerical results for the value of the cancellation option; the differences in the value of the cancellation option are uniformly small. In fact, most of the differences are less than one cent per \$100 notional amount. The bid-ask spread on these complex derivatives is likely at least an order of magnitude greater than the size of these differences. As before, the differences are both positive and negative in sign. The numerical values of the cancelable and noncancelable swaps do differ slightly between the finite difference and LSM techniques. In general, the values implied by the finite difference algorithm for the cancelable and noncancelable swaps are about four to seven cents higher than the corresponding values implied by the LSM approach. This systematic pattern is due to slight differences in the way that the two techniques discretize the continuous coupon payments and the continuous amortization feature. These differences produce minor differences in the levels of the swap values, but have almost no effect on the value of the cancellation option.

6. Jump-Diffusions and American Option Valuation

In this section, we illustrate how the LSM approach can be applied to value American options when the underlying asset follows a jump-diffusion process. In particular, we revisit the American put option considered in Section 3. To simplify the illustration, we focus on the basic jump-to-ruin model presented in Merton (1976). In this model, the stock price follows a geometric Brownian motion as in Equation (7) until a Poisson event occurs, at which point the stock price becomes zero. The dynamics for this jump-diffusion process are given by

$$dS = (r + \lambda)S\,dt + \sigma S\,dZ - S\,dq, \tag{16}$$

where q is an independent Poisson process with intensity λ. When a Poisson event occurs, the value of q jumps from zero to one, implying $dq = 1$, and the stock price jumps downward from S to zero. As in Merton, we assume that jump risk is nonsystematic and unpriced by the market. This assumption could easily be relaxed. Similarly, the LSM approach can be readily applied using much more complex jump-diffusion processes than in this example or the other examples given in Merton.

---

[16] Again, adding more basis functions has little effect on the value of the option. This provides numerical evidence that the number of basis functions needed to obtain effective convergence grows at a much slower rate than exponential as the dimensionality of the problem increases, consistent with Judd (1998).

Merton (1976) provides a closed-form solution for the value of a European option on the stock when its price follows the jump-to-ruin process in Equation (16). He also shows, however, that the price of an American option is given by a complex mixed differential-difference equation which is difficult to solve.[17]

To put the results into perspective, we compare the price and early exercise boundary for the American put option for the cases where there is no possibility of a jump $\lambda = .00$ and when a jump can occur with intensity $\lambda = .05$. Note that when $\lambda > 0$, the stock price process has a mass point at zero, and the distribution of the stock price is no longer conditionally log-normal. Furthermore, as λ increases, the conditional variance of the future stock price increases. Specifically, the variance of the stock price is

$$S^2(0) \exp(2r)\big(\exp((\lambda + \sigma^2)T) - 1\big). \tag{17}$$

To make the comparison more meaningful, we adjust the parameters in the two cases so that the means and variances are equal; the two cases differ in the shape of their conditional distributions but not in terms of their first two moments. From Equation (16), the mean of the risk-neutral distribution for the stock price is $S(0)\exp(rT)$ and is the same across cases because of the martingale restriction implied by the risk-neutral framework. To equalize variances, we assume that when $\lambda = 0$, $\sigma^2 = .09$. Similarly, when $\lambda = .05$, $\sigma^2 = .04$. With these parameter values, the two distributions for the stock price have the same means and variances. We use 26 exercise points per year in the LSM algorithm and use the same basis functions in these examples as in Section 3. We focus on the case $T = 1$.

Applying the LSM algorithm, the American put values are 3.80 for the $\lambda = .00$ case and 3.40 for the $\lambda = .05$ case. The European put values are 3.58 for the $\lambda = .00$ case and 3.23 for the $\lambda = .05$ case. The early exercise values are .22 and .17, respectively. Thus the values of the options are lower when there is a possibility of a jump, holding fixed the variance across the examples. This is intuitive since the diffusion coefficient in the $\lambda = .05$ case is only .20, while the diffusion coefficient in the $\lambda = .00$ case is .30. This means that in the absence of a jump, the option is less likely to be deep in the money in the $\lambda = .05$ case. If a jump occurs, of course, then the option is much more valuable than it would be otherwise. The results indicate, however, that the windfall gain to the optionholder from a jump does not offset the effects of the lower diffusion coefficient.

Since the value of the early exercise premium is less in the case where $\lambda = .05$, there is less incentive to keep the option alive. Consequently, it is not surprising that the optimal early exercise strategy is more aggressive

---

[17] Pham (1995) shows that the price of an American option can also be represented as the solution of a parabolic integrodifferential free boundary problem when the underlying price process exhibits jumps.
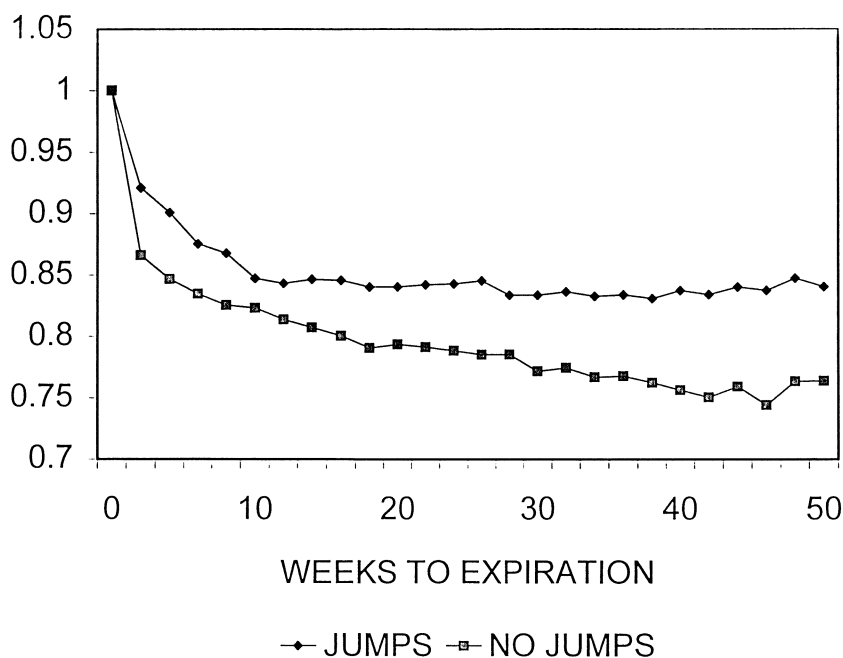
Figure 1
Graph of the early exercise boundary for an American put option
The early exercise boundary is shown as a proportion of the exercise price of the option. The jump graph shows the early exercise boundary when the underlying stock price follows a jump diffusion. The no-jump graph shows the early exercise boundary when the underlying stock price follows a pure diffusion process.

in the case where $\lambda = .05$ than in the case where $\lambda = .00$. To see this, Figure 1 plots the early exercise boundaries for the two cases. The early exercise boundaries are obtained by solving for the critical stock price at each exercise point at which the estimated conditional expectation function equals the immediate exercise value of the option. As shown, the early exercise boundary for the case where the stock price can jump is significantly higher than for the continuous case.

7. Valuing Swaptions in a String Model

To illustrate its generality, we apply the LSM approach to a deferred American swaption in a 20-factor string model. Swaptions are one of the most important and widely used derivatives in fixed-income markets. We focus on a basic swaption where the optionholder has the right to enter into a swap in which the optionholder receives fixed coupons and pays floating coupons on a semiannual cycle. The floating coupon paid at the end of the semiannual cycle is tied to the six-month rate determined at the beginning of the semi-annual cycle; the floating leg sets in advance and pays in arrears. Both legs of the swap are paid on an actual/actual basis. The swaption can only be

exercised on semiannual coupon payment dates and only after exchanging the coupons due on the payment date. As is typical, the swaption cannot be exercised until after a specific lockout period.

To provide a specific example, we focus on a 10 NC 1 swaption. This notation indicates the underlying swap has a final maturity of 10 years. The NC 1 (noncall 1) feature indicates that the swaption cannot be exercised until one year has elapsed; the swaption cannot be exercised until the second semiannual coupon payment date. Since there are 20 semiannual coupon payment dates during the life of the underlying 10-year swap, there are 18 possible exercise dates for the swaption; the swaption cannot be exercised at the first coupon payment date, and the swaption has no value at the final coupon payment date.

String models of the term structure have recently received a significant amount of attention in fixed-income markets. In these models, each point along the term structure is a separate random variable, where the term structure is defined either as a discount function or spot curve as in Ho and Lee (1986), a forward curve as in Heath, Jarrow, and Morton (1992), or as a par curve as in Longstaff, Santa-Clara, and Schwartz (1999). Important examples of string models include the recent articles by Goldstein (1997), Santa-Clara and Sornette (2001), Longstaff, Santa-Clara, and Schwartz (1999, 2000).

Since the underlying swap makes coupon payments at 20 different points in time, its value is sensitive to 20 different points along the curve. We implement a simple string model by assuming that each of these 20 points represents a separate but correlated factor, and model the joint dynamics of these 20 factors. Specifically, let $D(t, T)$ denote the value at time t of a zero-coupon bond with final maturity date T, where $t \leq T$. Since the expected rate of return on all securities must equal the riskless rate in the no-arbitrage equivalent martingale-measure framework, we can represent the no-arbitrage dynamics of the zero-coupon bond price by the following.

$$dD(t, T) = r(t)D(t, T)dt + \sigma(T - t)D(t, T)dZ_T, \qquad (18)$$

where $r(t)$ is the riskless rate, $\sigma(T - t)$ is a time-homogeneous volatility function, and Z_T is a standard Brownian motion specific to the zero-coupon bond with final maturity date T.

To operationalize this string model, we assume that the 20 factors are the 20 zero-coupon bond prices with maturities corresponding to the 20 coupon payment dates; specifically, $D(t, .5), D(t, 1.0), D(t, 1.5), \ldots, D(t, 10.0)$. We assume that the volatility function $\sigma(T-t)$ is piecewise constant; $\sigma(T-t)$ is constant over $.5(N - 1) < T - t \leq .5N$, $N = 1, 2, \ldots, 20$. For simplicity, we set $\sigma(T - t) = 0$ for values of $T - t < .5$. The remaining 19 values of $\sigma(T - t)$ are selected to approximate a cap volatility curve. Similarly, we assume that $r(t)$ is piecewise constant over six-month intervals;

we set $r(t)$ equal to the six-month rate defined by $-2\ln(D(t, t + 1/2))$. This discretization results in little loss of accuracy and guarantees that the price of a zero-coupon bond converges to one at its maturity date.[18]

The dynamics of the term structure are simulated by first solving the stochastic differential equation in Equation (18),

$$D(t + 1/2, T) = D(t, T) \exp\big(r(t)/2 - \sigma^2(T - t)/4$$
$$+ \sigma(T - t)\big(Z_T(t + 1/2) - Z_T(t)\big)\big). \qquad (19)$$

With this closed-form expression, the evolution of the term structure can be simulated over six-month periods. The only remaining issue is the correlation structure of the fundamental Brownian motions Z_T. To model the correlation matrix in a parsimonious way, we assume that the correlation between Z_i and Z_j, $i, j \leq T$ is given by the function $\rho_{ij} = \exp(-\kappa \mid i - j \mid)$, where $\kappa = .02$. This results in correlations among spot rates similar to those observed empirically. Alternatively, spot-rate correlations could be estimated using historical data and then directly incorporated into the simulation.

The simulation consists of paths where at each coupon date, the entire vector of zero-coupon bond prices is specified. From these zero-coupon bonds, the value of the underlying swap at that coupon date is easily computed by discounting the remaining fixed coupon payments; recall that the floating leg of the swap can be assumed to be worth par on coupon payment dates. Given the value of the swap, the basis functions are chosen to be a constant, the first three powers of the value of the underlying swap, and all unmatured discount bond prices with final maturity dates up to and including the final maturity date of the swap. Thus there are up to 22 basis functions in the regression. This specification results in values very similar to those obtained by using additional basis functions.

Table 5 reports the estimated values of the deferred American swaption, the corresponding European swaption, and the probabilities of early exercise at each coupon payment date for a variety of fixed coupon rates. As shown, the deferred American exercise feature has significant value; when the coupon rate is .0575, the deferred American swaption is more than three times as valuable as its European counterpart. This underscores the fact that the deferred American and European swaptions are fundamentally different derivatives despite their superficial similarities.

The coupon rate on the fixed leg of the swap also has a major effect on the properties of the swaptions. Clearly, the higher the coupon, the more valuable it is to enter the swap and receive the fixed coupons. For the European swaption, this translates into a higher probability of exercise at the exercise date.

---

[18] The minor discretization error induced by using the six-month rate as a proxy for $r(t)$ can easily be avoided by using more points along the term structure. For example, we could use daily, monthly, or even quarterly rates. Alternatively, we could simply develop the model in a discrete rather than continuous framework as is commonly done in practice.

Table 5

Exercise type	European	American	European	American	European	American
Coupon	.0575	.0575	.0605	.0605	.0635	.0635
Value	.739	2.577	1.529	3.278	2.711	4.204
Standard error	.011	.018	.016	.020	.021	.022
Ex. prob. 1	0.00	0.00	0.00	0.00	0.00	0.00
Ex. prob. 2	29.77	2.72	49.17	7.17	68.63	17.30
Ex. prob. 3	0.00	5.14	0.00	8.10	0.00	9.19
Ex. prob. 4	0.00	4.13	0.00	5.94	0.00	6.29
Ex. prob. 5	0.00	3.83	0.00	4.85	0.00	5.51
Ex. prob. 6	0.00	2.77	0.00	3.44	0.00	4.17
Ex. prob. 7	0.00	3.46	0.00	4.25	0.00	3.49
Ex. prob. 8	0.00	2.90	0.00	3.21	0.00	3.34
Ex. prob. 9	0.00	3.43	0.00	3.40	0.00	3.45
Ex. prob. 10	0.00	3.13	0.00	3.21	0.00	2.86
Ex. prob. 11	0.00	3.35	0.00	3.02	0.00	2.66
Ex. prob. 12	0.00	3.11	0.00	2.63	0.00	2.26
Ex. prob. 13	0.00	3.19	0.00	2.91	0.00	2.61
Ex. prob. 14	0.00	3.20	0.00	2.63	0.00	2.00
Ex. prob. 15	0.00	3.80	0.00	3.22	0.00	2.92
Ex. prob. 16	0.00	3.36	0.00	3.24	0.00	2.58
Ex. prob. 17	0.00	4.39	0.00	4.22	0.00	2.79
Ex. prob. 18	0.00	5.69	0.00	4.65	0.00	3.75
Ex. prob. 19	0.00	8.33	0.00	7.13	0.00	5.45
Ex. prob. 20	0.00	0.00	0.00	0.00	0.00	0.00
Total prob.	29.77	69.93	49.17	77.22	68.63	82.62

Estimated values and exercise probabilities for a deferred American swaption implied by the 20-factor string model. This swaption gives the optionholder the right to enter into a swap with a final maturity of 10 years and receive a fixed coupon and pay the six-month rate semiannually. The swaption is not exercisable until one year from the valuation date. The exercise probabilities are shown for each of the 20 semiannual coupon payment dates and represent the percentage of paths for which exercise occurred on that coupon payment date. Values are given per $100 notional amount. The simulation results are based on 20,000 paths.

In contrast, a higher fixed coupon does not necessarily imply a higher probability of exercise at a specific coupon date for the deferred American swaption. To see this, note that the probability of exercise at the 19th coupon date is 8.33% when the coupon is .0575, but is only 5.45% when the coupon is .0635. The total probability of exercise, however, is monotonic in the coupon rate. These results illustrate that the term structure of exercise probabilities for American options can display very complex patterns in a multifactor framework.

Since each point on the term structure affects the values of the deferred American and European swaptions, it is also interesting to compare the sensitivities of each swaption to each point on the curve. This is done by varying each of the six-month forward rates implied by the initial zero-coupon curve to express the risk exposures in a forward-space metric. As shown in Table 6, deferred American and European swaptions have major differences in their sensitivities to movements in the term structure. For example, the European swaption has the greatest sensitivity to the third forward. In contrast, the deferred American swaption has its greatest sensitivity to the 20th forward. These results illustrate the importance of incorporating the multifactor nature of the term structure in fixed-income derivative risk management.

Table 6

Forward rate	European	American
.0–.5	−.00008	−.00016
.5–1.0	−.00008	−.00016
1.0–1.5	−.00236	−.00048
1.5–2.0	−.00230	−.00077
2.0–2.5	−.00223	−.00089
2.5–3.0	−.00217	−.00117
3.0–3.5	−.00211	−.00134
3.5–4.0	−.00205	−.00142
4.0–4.5	−.00199	−.00151
4.5–5.0	−.00193	−.00156
5.0–5.5	−.00188	−.00148
5.5–6.0	−.00182	−.00143
6.0–6.5	−.00177	−.00176
6.5–7.0	−.00172	−.00181
7.0–7.5	−.00167	−.00187
7.5–8.0	−.00162	−.00180
8.0–8.5	−.00157	−.00195
8.5–9.0	−.00153	−.00186
9.0–9.5	−.00148	−.00199
9.5–10.0	−.00144	−.00218
Parallel shift	−.03380	−.02759

Sensitivity of swaption values to individual forward rates in the 20-factor string model. These sensitivities are computed by varying the indicated six-month forward rate while holding the others fixed. The American swaption gives the optionholder the right to enter into a swap with a final maturity of 10 years and receive a fixed coupon of .0605 and pay the six-month rate semiannually. The swaption is not exercisable until one year from the valuation date. The European swaption is the counterpart of the American swaption with the restriction that the option can only be exercised one year from the valuation date. The sensitivities shown are with respect to a one-basis point move in the forward per \$100 notional amount. The simulation results are based on 20,000 paths.

8. Numerical and Implementation Issues

In this section, we discuss a number of numerical and implementation issues associated with the LSM algorithm. These are discussed individually below.

8.1 Higher-dimensional problems

The numerical examples in Sections 3–5 benchmark the performance of the LSM algorithm for several low-dimensional problems which can be solved by standard finite difference techniques. As an additional benchmark, we also investigate the performance of the algorithm for a higher dimensional problem studied by Broadie and Glasserman (1997c), the valuation of an American options on the maximum of five risky assets.

In their article, Broadie and Glasserman (1997) apply a stochastic mesh approach to place bounds on the value an American call option on the maximum of five assets, where each asset has a return volatility of 20% and each return is independent of the others. The option has a three-year life and is exercisable three times per year. The assets each pay a 10% proportional dividend and the riskless rate is assumed to be 5%. The strike price of the option is 100, and the initial values of all assets are assumed to be the same and equal to either 90, 100, or 110. Using their algorithm, they are able to estimate a 90% confidence band for the value of the option. From Table 6 of

their article, the tightest 90% confidence bands reported are [16.602, 16.710], [26.101, 26.211], and [36.719, 36.842] for the cases where the initial asset values are 90, 100, and 110, respectively. Broadie and Glasserman report that computing these bounds requires slightly more than 20 hours apiece using a 266 MHz Pentium II processor.

We value this American call option using essentially the same simulation procedure used in Section 3. Specifically, we use 50,000 paths and choose 19 basis functions consisting of a constant, the first five Hermite polynomials in the maximum of the values of the five assets, the four values and squares of the values of the second through fifth highest asset prices, the product of the highest and second highest, second highest and third highest, etc., and finally, the product of all five asset values. The values of the American call option given by the LSM algorithm are 16.657, 26.182, and 36.812 for the cases where the initial asset values are 90, 100, and 110, respectively. In each of these cases, the LSM value is within the tightest bounds given by the Broadie and Glasserman algorithm. We note that computing these values by LSM requires only one to two minutes apiece using a 300 MHz Pentium II processor.

8.2 Least squares

In this article, we use ordinary least squares to estimate the conditional expectation function. In some cases, however, it may be more efficient to use other techniques such as weighted least squares, generalized least squares, or even GMM in estimating the conditional expectation function. For example, if the process for the state variables has state dependent volatility, the residuals from the regression may be heteroskedastic and these alternative least squares techniques may have advantages.

In estimating the least squares regressions, it may be noted that the R^2s from the regressions are often somewhat low. The reason for this is simply the volatility of realized cash flows around their conditional expected values. The R^2 from the regression measures the percentage of the total variation in the ex post cash flows explained by the variation in the conditional expectation function; a low R^2 simply means that the volatility of unexpected cash flows is large relative to the volatility of expected cash flows. Thus low R^2 are to be expected when unexpected cash flows are highly volatile. In general, since the LSM algorithm is based on conditional first moments rather than second moments, the R^2s from the regression should have little impact on the quality of the LSM approximation to the American option value.

8.3 Choice of basis functions

Extensive numerical tests indicate that the results from the LSM algorithm are remarkably robust to the choice of basis functions. For example, we use the first three Laguerre polynomials as basis functions in the American put illustration in Section 3. We obtain results that are virtually identical to those

reported in Table 1 when we use S, S^2, and S^3 as basis functions, when we use the first three Hermite polynomials as basis functions, or when we use three trigonometric functions as basis functions. Similarly for all of the other examples presented in the article. As reported earlier, few basis functions are needed to closely approximate the conditional expectation function over the relevant range where early exercise may be optimal.

While the results are robust to the choice of basis functions, it is important to be aware of the numerical implications of the choice. For example, the weighted Laguerre polynomials used in the American put illustration in Section 3 include an exponential term in the stock price S. In Table 1, however, the stock price ranges from 36 to 44. Thus, directly applying the weighted Laguerre polynomials to the problem could result in computational underflows. To avoid this problem, we renormalized the American put example by dividing all cash flows and prices by the strike price, and estimating the conditional expectation function in the renormalized space; the results reported in Table 1 are based on this renormalization. Note that this is only for numerical convenience; the option value is unaffected since we discount back the unnormalized value of the cash flows along each path to obtain its value. We recommend normalizing appropriately to avoid numerical errors resulting from scaling problems.

Finally, the choice of basis functions also has implications for the statistical significance of individual basis functions in the regression. In particular, some choices of basis functions are highly correlated with each other, resulting in estimation difficulties for individual regression coefficients akin to the multicolinearity problem in econometrics. This difficulty does not affect the LSM algorithm since the focus is on the fitted value of the regression rather than on individual coefficients; the fitted value of the regression is unaffected by the degree of correlation among the explanatory variables. However, if the choice of basis functions leads to a cross-moment matrix that is nearly singular, then numerical inaccuracies in some least squares algorithms may affect the functional form of estimated conditional expectation function. To avoid these types of numerical problems, the regressions are estimated using the double-precision DLSBRR algorithm in IMSL which estimates least squares via an iterative-refinement algorithm. We also cross-checked the results by estimating the regressions using a variety of alternative procedures such as Cholesky-decomposition and QR-algorithm least squares techniques.

8.4 Computational speed

An important advantage of simulation techniques is that they lend themselves well to parallel computing architecture. For example, we could generate 5,000 paths using a single CPU, or we could generate 100 paths each on 50 CPUs. In many large-scale applications, computational speed is far more important than the cost of hardware; valuation and risk management by simulation is ideally suited for these applications. Furthermore, for some

types of large-scale applications such as valuing large portfolios of fixed-income or mortgage derivatives, the same paths can be used for all options; significant computational efficiencies are obtained by only having to generate paths once. From the perspective of the LSM algorithm, the only constraint on parallel computation is that the regression needs to use the cross-sectional information in the simulation. Given the speed at which regressions can be estimated, however, this bottleneck involves little loss of computational efficiency. Furthermore, there are many ways in which regressions could be estimated using individual CPUs, and then aggregated across CPUs to form a composite estimate of the conditional expectation function. Finally, the use of quasi-Monte-Carlo techniques in conjunction with the LSM algorithm may lead to significant improvements in computational speed and efficiency. Important recent examples of the application of quasi-Monte-Carlo techniques include Morokoff and Caflisch (1994, 1995).

9. Conclusion

This article presents a simple new technique for approximating the value of American-style options by simulation. This approach is intuitive, accurate, easy to apply, and computationally efficient. We illustrate this technique using a number of realistic examples, including the valuation of an American put when the underlying stock price follows a jump-diffusion as well as the valuation of a deferred American swaption in a 20-factor string model of the term structure.

As a framework for valuing and risk managing derivatives, simulation has many important advantages. With the ability to value American options, the applicability of simulation techniques becomes much broader and more promising, particularly in markets with multiple factors. Furthermore, simulation techniques make it much easier to implement advanced models such as Heath, Jarrow, and Morton (1992) or Santa-Clara and Sornette (1997) in practice.

Appendix

Proof of Proposition 1. By definition, the value of the underlying asset X_{t_k} is $\mathcal{F}_{t_k}$-measurable. Similarly, the immediate exercise value of the option is $\mathcal{F}_{t_k}$-measurable. By construction, $F_M(\omega; t_k)$ is a linear function of $\mathcal{F}_{t_k}$-measurable functions of X_{t_K}, and is therefore $\mathcal{F}_{t_k}$-measurable. Hence the event that the immediate exercise value is greater than zero and greater than or equal to $F_M(\omega; t_k)$ is in $\mathcal{F}_{t_k}$, and therefore, the LSM rule to exercise when this event occurs defines a stopping time. Denote the present value of following this stopping rule by E_θ. Since $V(X)$ is the supremum of the present values obtained over the set of all stopping times, $V(X) \geq E_\theta$. Since the functional form of $F_M(\omega; t_k)$ is the same across all paths, the discounted cash flows $\mathrm{LSM}(\omega_i; M, K)$ are independently and identically distributed, and the strong law of large numbers [Billingsley (1979; Theorem 6.1)] implies that

$$\Pr\left[\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} LSM(\omega_i; M, K) = E_\theta \right] = 1.$$

This result, combined with the inequality $V(X) \geq E_\theta$, implies the result. ∎

Proof of Proposition 2. At time t_2, the LSM stopping strategy is the same as the optimal strategy; the option is exercised if it is in the money. Under the given assumptions, the conditional expectation function $F(\omega; t_1)$ is a function only of X_{t_1}. If $F(\omega; t_1)$ satisfies the indicated conditions, then Theorem IV.9.1 of Sansone (1959) implies that the convergence of $F_M(\omega; t_1)$ to $F(\omega; t_1)$ is uniform in M on the set $(0, \infty)$, where the first M Laguerre polynomials are used as the set of basis functions. This implies that for a given ϵ, there exists an M such that $\sup_{X_{t_1}} | F(\omega; t_1) - F_M(\omega; t_1) | < \epsilon/2$. From the integrability conditions and Theorem 3.5 of White (1984), the fitted value of the LSM regression $\widehat{F}_M(\omega; t_1)$ converges in probability to $F_M(\omega; t_1)$ as $N \to \infty$,

$$\lim_{N \to \infty} \Pr\big[| F_M(\omega; t_1) - \widehat{F}_M(\omega; t_1) | > \epsilon/2\big] = 0.$$

Thus, for any ϵ, there is an M such that

$$\lim_{N \to \infty} \Pr\big[| F(\omega; t_1) - \widehat{F}_M(\omega; t_1) | > \epsilon\big] = 0.$$

To complete the proof, we partition Ω into five sets: 1) the set of paths where the option is exercised at time t_1 under both the optimal and the LSM strategy, 2) the set of paths where the option is not exercised at time t_1 under either the optimal or LSM strategies, 3) the set of paths where the option is exercised at time t_1 under the LSM strategy but not under the optimal strategy, 4) the set of paths where the option is exercised at time t_1 under the optimal strategy, but not under the LSM strategy, and 5) a zero-probability set of paths for which the difference between $F(\omega; t_1)$ and $\widehat{F}_M(\omega; t_1)$ is greater than ϵ as $N \to \infty$. Now consider a portfolio consisting of a long position in an option exercised using the LSM strategy, an investment of ϵ in a money market account, and a short position in an option exercised using the optimal strategy. For paths in set 3), at time t_1, use the funds from the money market account and the option exercise to purchase a European option with final maturity date t_2. For paths in set 4), at time t_1, use the funds from the money market account and from shorting a European option with final maturity date t_2 to fund the cash outflow from the option exercise. It is easily shown that this strategy results in cash flows that are greater than or equal to zero for each path in sets 1), 2), 3), and 4). Since the pathwise cash flows are nonnegative, averages over paths are nonnegative, and the result follows from a standard no-arbitrage argument, the definition of $V(X)$, and the law of large numbers. ∎

References

Abramowitz, M., and I. A. Stegun, 1970, *Handbook of Mathematical Functions*, Dover Publications, New York.

Amemiya, T., 1985, *Advanced Econometrics*, Basil Blackwell, London, UK.

Averbukh, V., 1997, "Pricing American Options Using Monte Carlo Simulation," Ph.D. dissertation, Cornell University.

Barraquand, J., and D. Martineau, 1995, "Numerical Valuation of High Dimensional Multivariate American Securities," *Journal of Financial and Quantitative Analysis*, 30, 383–405.

Barron, A., 1993, "Universal Approximation Bounds for Superpositions of a Sigmoidal Function," *IEEE Transactions, Information Theory*, 39, 930–945.

Bensoussan, A., 1984, "On the Theory of Option Pricing," *Acta Applicandae Mathematicae*, 2, 139–158.

Billingsley, P., 1979, *Probability and Measure*, Wiley, New York.

Black, F., and M. Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637–654.

Bossaerts, P., 1989, "Simulation Estimators of Optimal Early Exercise," working paper, Carnegie-Mellon University.

Broadie, M., J. Detemple, E. Ghysels, and O. Torres, 1998, "American Options with Stochastic Dividends and Volatility: A Nonparametric Investigation," working paper, Columbia University.

Broadie, M., and P. Glasserman, 1997a, "Monte Carlo Methods for Pricing High-Dimensional American Options: An Overview," working paper, Columbia University.

Broadie, M., and P. Glasserman, 1997b, "Pricing American-Style Securities Using Simulation," *Journal of Economic Dynamics and Control*, 21, 1323–1352.

Broadie, M., and P. Glasserman, 1997c, "A Stochastic Mesh Method for Pricing High-Dimensional American Options," working paper, Columbia University.

Broadie, M., P. Glasserman, and G. Jain, 1997, "Enhanced Monte Carlo Estimates for American Option Prices," *Journal of Derivatives*, 5, 25–44.

Carriere, J., 1996, "Valuation of Early-Exercise Price of Options Using Simulations and Nonparametric Regression," *Insurance: Mathematics and Economics*, 19, 19–30.

Carr, P., 1998, "Randomization and the American Put," *Review of Financial Studies*, 11, 597–626.

Cox, J., J. E. Ingersoll, and S. A. Ross, 1985, "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica*, 53, 363–384.

Cox, J., and S. A. Ross, 1976, "The Valuation of Options for Alternative Stochastic Processes," *Journal of Financial Economics*, 3, 145–166.

Duffie, D., 1996, *Dynamic Asset Pricing Theory*, Princeton University Press, Princeton, N.J.

Garcia, D., 1999, "A Monte Carlo Method for Pricing American Options," working paper, University of California, Berkeley.

Goldstein, R., 1997, "The Term Structure of Interest Rates as a Random Field," working paper, Ohio State University.

Harrison, J. M., and D. M. Kreps, 1979, "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory*, 20, 381–408.

Harrison, J. M., and S. R. Pliska, 1981, "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Processes and their Applications*, 11, 261–271.

Heath, D., R. Jarrow, and A. Morton, 1992, "Bond Pricing and the Term Structure of Interest Rates," *Econometrica*, 60, 77–106.

Ho, T., and S. Lee, 1986, "Term Structure Movements and Pricing Interest Rate Contingent Claims," *Journal of Finance*, 41, 1011–1029.

Ibanez, A., and F. Zapatero, 1998, "Monte Carlo Valuation of American Options through Computation of the Optimal Exercise Frontier," working paper, University of Southern California.

Keane, M., and K. Wolpin, 1994, "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation: Monte Carlo Evidence," *Review of Economics and Statistics*, 76, 648–672.

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

Karatzas, I, 1988, "On the Pricing of American Options," *Applied Mathematics and Optimization*, 17, 37–60.

Lamberton, D., and B. Lapeyre, 1996, *Stochastic Calculus Applied to Finance*, Chapman & Hall, London, UK.

Longstaff, F., P. Santa-Clara, and E. Schwartz, 1999, "Throwing Away a Billion Dollars: The Cost of Suboptimal Exercise Strategies in the Swaptions Market," working paper, University of California, Los Angeles.

Longstaff, F., P. Santa-Clara, and E. Schwartz, 2000, "The Relative Valuation of Interest-Rate Caps and Swaptions: Theory and Empirical Evidence," working paper, University of California, Los Angeles.

Merton, R. C., 1973, "The Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science*, 4, 141–183.

Merton, R. C., 1976, "Option Pricing When Underlying Stock Returns Are Discontinuous," *Journal of Financial Economics*, 3, 125–144.

Morokoff, W. J., and R. E. Caflisch, 1994, "Quasi-Random Sequences and Their Discrepancies," *SIAM Journal of Computing*, 15, 1251–1279.

Morokoff, W. J., and R. E. Caflisch, 1995, "Quasi-Monte Carlo Integration," *Journal of Computational Physics*, 122, 218–230.

Pham, H., 1995, "Optimal Stopping, Free Boundary and American Option in a Jump Diffusion Model," working paper, Universite Paris IX Dauphine.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1992, *Numerical Recipes in Fortran*, 2nd ed., Cambridge University Press, Cambridge, UK.

Raymar, S., and M. Zwecher, 1997, "A Monte Carlo Valuation of American Call Options on the Maximum of Several Stocks," *Journal of Derivatives*, 5, 7–23.

Refenes, A., A. Burgess, and Y. Bentz, 1997, "Neural Networks in Financial Engineering: A Study in Methodology," *IEEE Transactions, Neural Networks*, 8, 1222–1267.

Royden, H. L., 1968, *Real Analysis*, MacMillan, New York.

Sansone, G., 1959, *Orthogonal Functions*, Interscience Publishers, New York.

Santa-Clara, P., and D. Sornette, 2001, "The Dynamics of the Forward Interest Rate Curve with Stochastic String Shocks," *The Review of Financial Studies*, 14, 149–185.

Tilley, J. A., 1993, "Valuing American Options in a Path Simulation Model," *Transactions of the Society of Actuaries*, 45, 83–104.

Tsitsiklis, J., and B. Van Roy, 1999, "Optimal Stopping of Markov Processes: Hilbert Space Theory, Approximation Algorithms, and an Application to Pricing High-Dimensional Financial Derivatives," *IEEE Transactions on Automatic Control*, 44, 1840–1851.

Vasicek, O., 1977, "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5, 177–188.

White, H., 1984, *Asymptotic Theory for Econometricians*, Academic Press, New York.