ELSEVIER Insurance: Mathematics and Economics 19 (1996) 19–30

INSURANCE MATHEMATICS & ECONOMICS

Valuation of the early-exercise price for options using simulations and nonparametric regression

Jacques F. Carriere

Faculty of Management, University of Manitoba, Winnipeg, Man., Canada R3T 5V4

Received June 1994; revised February 1996

Abstract

This article shows how to value the optimal stopping time for any Markovian process in finite discrete time. Specifically, the article focuses on the valuation of American options using simulations of stochastic processes. It also shows that the estimation of the decision rule to exercise early is equivalent to the estimation of a series of conditional expectations. For Markov processes, these conditional expectations can be estimated with nonparametric regression techniques. This article shows how to approximate the conditional expectations and the resulting early-exercise decision rule with spline and local regression.

Keywords: American options; Markov processes; Stopping times; Arbitrage-free pricing; Martingales; Splines; Locally weighted regression

1. Introduction

This article shows how to value the early exercise privilege for American options in finite discrete time. In fact, the technique that we present can be used to approximate the optimal stopping for any Markovian process in discrete time. For an introduction to options, consult Boyle (1992).

Using the theory of stopping times, we present a backward induction algorithm that allows us to calculate the value of the early exercise privilege. This algorithm is based on the successive calculation of conditional expectations that are usually difficult to evaluate explicitly. We suggest that these conditional expectations can be approximated with a nonparametric regression on simulated data. The paper also compares our method to the one given in Tilley (1993).

The article by Tilley (1993) dispels the belief that American options cannot be valued efficiently through simulation. Tilley's article presents a general algorithm for estimating the value of American options in an arbitrage-free setting.

A weakness of the Tilley algorithm is the biased nature of the estimator. This paper shows why the bias exists but more importantly it shows how to construct an unbiased estimator. Next, Tilley's article provides scant justification for the form of the algorithm. In contrast, this paper uses the theory of optimal stopping to justify the form of our algorithm which is a sequential regression algorithm.

PII S0167-6687(96)00004-2

2. The intrinsic value of an option

In this section, we define the intrinsic values of various options and we present some notation and basic definitions. We assume that investors can buy or sell risk-free bonds and risky assets like commodities, stocks and derivatives. We assume that the number of commodities, stocks and derivatives is finite. All commodities, stocks and derivatives are tradable and their prices are random. Investors are allowed to hold any portfolio of securities with no budget constraints. This means that an investor, with no money, can borrow from the bond market to finance investments in the stock market. We also assume that investors are risk-neutral. This means that if the price for two securities are the same, then an investor is indifferent to them.

Our asset-pricing model is a model in finite discrete time which means that assets can only be bought and sold at the times $n = 0, 1, \ldots, N$. Let $1 \geq v_n \geq 0$ denote the price of a risk-free bond at time 0 that matures at time n at a unit redemption value. Usually $v_0 = 1$ and $v_n > v_{n+1}$. The price of an individual commodity or stock at time n is denoted as $S_n \geq 0$ while the value of the early exercise option for an individual derivative at time 0 is denoted as $V_0 \geq 0$. We assume that all investors have access to information in the form of the observed variables $\boldsymbol{H}_n$. The information contained in $\boldsymbol{H}_n$ includes all the prices of tradable securities at time n and other data as well. The past history of these observed variables is denoted as $\mathcal{H}_n$ and we assume that $\mathcal{H}_n$ is equal to the sub-σ-field $\mathcal{H}_n \equiv \sigma\{\boldsymbol{H}_k | 0 \leq k \leq n\}$.

The main focus of this paper is the problem of valuing the early exercise privilege of an option in finite discrete time. Let $I_n \equiv I_n(\boldsymbol{H}_n)$ denote the *intrinsic* value of the option at time n. The intrinsic value I_n is the amount that the investor would get at time n if the investor exercised the early exercise option. Let us give examples of the intrinsic value function. The American *call* option is a derivative security where the intrinsic value is $I_n \equiv \max\{0, S_n - K\}$ for $n = 0, 1, \ldots, N$. The value K denotes the strike price while S_n denotes the stock price. Another example of a derivative security is an European *put* option. In this case the intrinsic value function is equal to $I_n = 0$ if $n < N$ and $I_N \equiv \max\{0, K - S_N\}$. An example of an *exotic* American put option is one where the intrinsic value is

$$I_n \equiv \max\{0, K_n - \min(S_{n,1}, S_{n,2})\} = \max\{0, K_n - S_{n,1}, K_n - S_{n,2}\}$$

for $n = 0, 1, \ldots, N$. In this case $S_{n,1}$ and $S_{n,2}$ denote stock prices while K_n denotes a nonconstant strike price. This type of exotic option was investigated by Tang and Vetzal (1994). Generally, the intrinsic value of an option is any function of the observed variables $\boldsymbol{H}_n$.

3. The value of random cash flows under risk neutrality

In this section, we show how to value cash flows at time $n = 0$ when investors are risk neutral. The ideas in this section are well-known but are presented for completeness. Consider a portfolio of risk-free bonds where the number of bonds maturing at time $n = 0, 1, \ldots, N$ is fixed at $x_n \in \mathbb{R}$. In a perfect market, the price of this portfolio at time $n = 0$ is simply equal to

$$\sum_{n=0}^{N} v_n \times x_n. \tag{3.1}$$

Next, consider a random cash flow $\mathbf{X} = (X_0, \ldots, X_N)^T \in \mathbb{R}^N$ where $|E\{X_n|\mathcal{H}_0\}| < \infty$ for all n. In a risk-neutral market, investors are indifferent to the random cash-flow $\mathbf{X}$ and the sure cash flow $E(\mathbf{X}) = \{E(X_0|H_0), \ldots, E(X_N|\mathcal{H}_0)\}^T$, which implies that the price paid for $\mathbf{X}$ is the same as the price paid for $E(\mathbf{X})$. Therefore the price paid for $\mathbf{X}$ or $E(\mathbf{X})$ is simply equal to

$$P = \sum_{n=0}^{N} v_n E(X_n|\mathcal{H}_0), \tag{3.2}$$

because in a perfect market the investor can borrow P and purchase the sure cash-flow $E(X)$ by issuing $E(X_n|\mathcal{H}_0)$ bonds that mature at the times $n = 0, 1, \ldots, N$.

Consider the implications of (3.2) with respect to the term structure of interest rates. Let $v_{n,m}$ denote the price of a risk-free bond at time n that matures at time m at a unit redemption value. Consider the cash flow where $X_n = -v_{n,m}$ and $X_m = 1$. This represents a situation where the investor lends out an amount of $v_{n,m}$ at time n and is repaid an amount of 1 at time m. In out market, an investor always has the privilege of lending money at no cost. This means that the value or price of this transaction at time 0 is simply 0. Using (3.2), we find that $0 = -v_n \times E(v_{n,m}|\mathcal{H}_0) + v_m$. Therefore,

$$E(v_{n,m}|\mathcal{H}_0) = v_m / v_n. \tag{3.3}$$

In other words, our valuation method is consistent with the *Local Expectations Hypothesis* of the term structure of interest rates. The reader can consult Cox, et al. (1981) for an extensive discussion of this hypothesis. They show that it is the only term structure model that obtains in equilibrium for a market model in continuous time. They also have a discussion on the hypothesis of discrete time.

Let us discuss the implications for pricing the early exercise privilege of an option when a random exercise strategy is used. At any time, the holder of a derivative security can exercise the early exercise option. This decision must be made on the basis of the observed variables $\boldsymbol{H}_0, \ldots, \boldsymbol{H}_n$ and not the future variables $\boldsymbol{H}_{n+1}, \boldsymbol{H}_{n+2} \ldots$ This means that the exercise time is a stopping time, which is denoted as τ, and at the exercise time the cash flow to the holder of the derivative is I_τ. This means that the random cash flow at time n is equal to $I_n \times 1(\tau = n)$ which implies that the value (at $n = 0$) of the strategy based on $\tau \geq 0$ is

$$\begin{aligned}
&\sum_{n=0}^{N} v_n \, E\left\{I_n \times 1(\tau = n)|\mathcal{H}_0\right\} \\
&\qquad = E\left\{\sum_{n=0}^{N} v_n \times I_n 1(\tau = n)|\mathcal{H}_0\right\} \\
&\qquad = E\{v_\tau \times I_\tau|\mathcal{H}_0\}.
\end{aligned} \tag{3.4}$$

This valuation formula assumes that v_n is not random because the information set $\mathcal{H}_0$ contains this information. Tilley (1993) presents a more general form of (3.4), where v_n is random.

4. Valuation using an optimal stopping time

In this section, we give a backward induction algorithm for optimal stopping in finite discrete time and we relate it to the valuation of the early exercise option for derivative securities.

Let $n = 0, 1, \ldots, N$ and let $\boldsymbol{H}_n$ denote the observed variables at time n. Suppose that a reward $R_n \equiv R_n(\boldsymbol{H}_n)$ is a function of the observed variables. In our application, the reward at time n is the discounted intrinsic value and so

$$R_n = v_n \times I_n. \tag{4.1}$$

Consider the history or sub-σ-field $\mathcal{H}_n \equiv \sigma\{\boldsymbol{H}_k|0 \leq k \leq n\}$ and the stochastic sequence $\{R_n, \mathcal{H}_n\}_{n=0}^{N}$ where R_n is integrable. In this discussion, we focus on stopping times, τ, that assume values on the integers $0, 1, \ldots, N$. A stopping time is simply a random variable with the property that the event $\{\tau = n\}$ is in the set $\mathcal{H}_n$ for all n. Let C_n denote the class of all stopping times τ such that $Pr(n \leq \tau \leq N) = 1$. Next, let

$$\begin{aligned}
V_n &\equiv \sup\nolimits_{\tau \in C_n} E(R_\tau|\mathcal{H}_n), \\
R_\tau &= \sum_{k=n}^{N} R_k \times 1\{\tau = k\}.
\end{aligned} \tag{4.2}$$

Note that $V_n, n \geq 1$, is not the value of the early exercise option at time n because we are discounting to time 0. Next, define

$$\tau_n \equiv \sum_{k=n}^{N} k \times 1\{V_h > R_h \text{ for } n \leq h < k \text{ and } V_k = R_k\}. \tag{4.3}$$

The backward induction theorem from Chow et al. (1971) states that the stopping time τ_n is optimal. The backward induction theorem is as follows.

Theorem 1.
(i) $V_N = R_n$ *and if* $n = 0, 1, \ldots, N-1$ *then* $V_n = \max\{R_n, E(V_{n+1}|\mathcal{H}_n)\}$,
(ii) $\tau_n \in C_n$ *and* $V_n = E\{R_{\tau n}|\mathcal{H}_n\} \forall n = 0, 1, \ldots, N$,
(iii) $E(V_n|\mathcal{H}_m) = \sup_{\tau \in C_n} E(R_\tau|\mathcal{H}_m) \forall 0 \leq m < n \leq N$.

The proof of this result can be found in Chow et al. (1971, p. 50), a treatise about optimal stopping times. The backward algorithm is actually described in part (i) of theorem 1. Theorem 1 essentially states that the value of the early exercise option for a derivative is equal to

$$V_0 = E\{R_{\tau_0}|H_0\} = E\left\{\sum_{n=0}^{N} v_n \times I_n \times 1(\tau_0 = n)|\mathcal{H}_0\right\}, \tag{4.4}$$

where τ_0 is the first time $n = 0, 1, \ldots, N$ such that $v_n \times I_n \geq E(V_{n+1} \mid \mathcal{H}_n)$. In this representation, $V_{N+1} = 0$. Note that usually $E\{R_{\tau_0} \mid \mathcal{H}_0\} > E\{R_0|\mathcal{H}_0\} = R_0$ and so $V_0 = E(V_1|\mathcal{H}_0)$ and we never exercise at time $n = 0$.

Let us give an equivalent representation of the *backward algorithm*. This alternate way of finding the optimal stopping time was used by Tilley (1993). Define

$$V_n^* \equiv V_n/v_n \quad \text{and} \quad d_n \equiv v_{n+1}/v_n. \tag{4.5}$$

Then we must have $V_N^* = I_N$ and for $n = N-1, \ldots, 1, 0$, we must have

$$V_n^* = \max\{I_n, d_n E(V_{n+1}^*|\mathcal{H}_n)\}. \tag{4.6}$$

The optimal stopping time τ_0 can now be described as the first time $n = 0, 1, \ldots, N$ such that $I_n \geq d_n E(V_{n+1}^*|\mathcal{H}_n)$. In conclusion, we find that

$$V_0^* = V_0. \tag{4.7}$$

During our simulations, we found that the approximations of V_0 based on V_n^* were almost identical to those based on V_n. Differences between the two were attributed to roundoff errors during the approximation process.

4.1. *Valuation of European options*

For European options, the intrinsic value function is equal function is equal to $I_n = 0$ if $n < N$. But if $n = N$ then

$$I_N = \begin{cases} \max\{0, K - S_N\} & \text{(put option)}, \\ \max\{0, S_N - K\} & \text{(call option)}, \end{cases} \tag{4.8}$$

assuming a constant strike price. Using the algorithm in Theorem 1*(i)*, we immediately find that in all cases, $\tau_0 = N$, and so

$$V_0 = \begin{cases} v_N \times E\{\max\{0, K - S_N\}|\mathcal{H}_0\} & \text{(put option)}, \\ v_N \times E\{\max\{0, S_N - K\}|\mathcal{H}_0\} & \text{(call option)}. \end{cases} \tag{4.9}$$

4.2. Valuation of an American call option

In this section, we show that the formula for an American call option is the same as the formula for the European call option, shown in (4.9). This classical result obtains when we assume that the discounted price process $\{v_n \times S_n\}_{n=0}^N$ is a martingale adapted to the history $\mathcal{H}_n$. As Schachermayer (1992) shows, this is equivalent to arbitrage-free pricing. Assuming that we have a martingale, we find that $\tau_0 = N$. This is true because if $n = 0, 1, \ldots, N-1$, then

$$\begin{aligned}
&E(V_{n+1}|\mathcal{H}_n)\\
&= E(\max\{v_{n+1} \times I_{n+1}, E(V_{n+2}|\mathcal{H}_{n+1})\}|\mathcal{H}_n)\\
&\geq E(v_{n+1} \times I_{n+1}|\mathcal{H}_n)\\
&= E(v_{n+1} \times \max\{0, S_{n+1} - K\}|\mathcal{H}_n)\\
&> \max\{0, E(v_{n+1}S_{n+1}|\mathcal{H}_n) - v_{n+1}K\}\\
&= \max\{0, v_n S_n - v_{n+1}K\}\\
&\geq v_n \times I_n,
\end{aligned}$$

assuming that the conditional density of S_{n+1} given S_n is absolutely continuous and greater than 0 on the positive reals. In other words, $V_n > R_n$ for all $n = 0, 1, \ldots, N-1$. In the case of an American put option, the optimal stopping time usually cannot be described simply and so no explicit formula for V_0 exists. This is also true of other derivative securities like exotic options. We now show how V_0 can be approximated in these cases.

5. The sequential approximation algorithm

In this section, we present our simulation algorithm for calculating the value, V_0, of the early exercise option for derivative securities. Let $M = 1, 2, \ldots$ denote the total number of sample paths generated in the simulation study and let

$$\mathsf{X}_m \equiv (\boldsymbol{H}_{0,m}, \ldots, \boldsymbol{H}_{N,m}), \tag{5.1}$$

denote the mth sample path and let $\boldsymbol{H}_{n,m}$ denote the observed variables at time n from the mth replication of the process. We assume that $\mathsf{X}_1, \mathsf{X}_2, \ldots$ are independent and identically distributed random processes. These simulated data are used to approximate the conditional expectations $E(V_{n+1}|\mathcal{H}_n)$ for $n = 0, 1, \ldots, N-1$. Define $\boldsymbol{H}_n \equiv \boldsymbol{H}_{n,1}$. To simplify the analysis, we assume that the process $\{\boldsymbol{H}_n, \mathcal{H}_n\}_{n=0}^N$ is *Markovian*. This implies that $E(V_{n+1}|\mathcal{H}_n) = E(V_{n+1}|\boldsymbol{H}_n)$, which implies that $V_n \equiv V_n(\boldsymbol{H}_n)$ is a function of $\boldsymbol{H}_n$ because $R_n \equiv R_n(\boldsymbol{H}_n)$ and $V_n = \max\{R_n, E(V_{n+1}|\mathcal{H}_n)\}$.

Fix n and suppose that we know the function $V_{n+1}(\cdot)$. Theoretically, we can use *nonparametric regression analysis* to estimate $E(V_{n+1}(\boldsymbol{H}_{n+1})|\boldsymbol{H}_n)$ with the simulated variables $\{\boldsymbol{H}_{n,m}, V_{n+1}(\boldsymbol{H}_{n+1,m})\}, m = 1, \ldots, M$. A good reference for nonparametric regression is Hardle (1990). An extensive discussion of spline techniques, a type of nonparametric regression method, can be found in Seber and Wild (1989). The local regression techniques presented in Cleveland and Devlin (1988) may also be used for nonparametric regression. Obviously, we do not know $V_{n+1}(\cdot)$ and so it must also be approximated. This leads to a recursive algorithm. At this point, it may be instructive to note that the "bundling" algorithm presented by Tilley (1993) is actually a regression method using a crude kernel smoothing techniques. Let $\hat{V}_n$ denote the approximation of V_n. To approximate V_0, we simply calculate the sequence $\hat{V}_N, \ldots, \hat{V}_0$ Where $\hat{V}_N(\cdot) \equiv R_N(\cdot)$ and

$$\hat{V}_0 = \frac{1}{M}\sum_{m=1}^{M} \hat{V}_1(\boldsymbol{H}_{1,m}), \tag{5.2}$$

and the estimates $\hat{V}_1(\cdot), \ldots, \hat{V}_{N-1}(\cdot)$ are based on the regression function estimators. Using this sequence, we can generate M replications of the optimal stopping time as follows:

$$\hat{\tau}_{0,m} = \sum_{k=0}^{N} \times 1\left\{\hat{V}_h(\boldsymbol{H}_{h,m}) > R_h(\boldsymbol{H}_{h,m}), 0 \le h < k, \hat{V}_k(\boldsymbol{H}_{k,m}) = R_k(\boldsymbol{H}_{k,m})\right\}. \tag{5.3}$$

Define $I_{n,m} \equiv I_n(\boldsymbol{H}_{n,m})$ and $z_{n,m} \equiv 1\{\hat{\tau}_{0,m} = n\}$. Using the empirical analog of (4.4), we get another way of estimating V_0. This is

$$\tilde{V}_0 = \frac{1}{M}\sum_{m=1}^{M}\sum_{n=0}^{N} v_n \times I_{n,m} \times z_{n,m}. \tag{5.4}$$

Tilley (1993) recommends estimators of the form $\tilde{V}_0$ to approximate V_0 and so do we. Later, we will compare the estimators $\tilde{V}_0$ and $\hat{V}_0$ and find that $\hat{V}_0$ is biased while $\tilde{V}_0$ is less biased than $\hat{V}_0$.

5.1. Calculating unbiased estimates

The statistic $\hat{V}_n$ will be biased for all $n < N$. It is sufficient to prove this when $n = N - 1$. Let $\hat{H}_n$ denote the estimate of $E(V_{n+1}|\mathcal{H}_n)$ and suppose that $\hat{H}_{n-1}$ is an unbiased estimator of $E(V_N|\mathcal{H}_{N-1}) = E(R_N|\mathcal{H}_{N-1})$, then

$$\begin{aligned} &E\ (\ \hat{V}_{N-1}|\mathcal{H}_{N-1}) \\ &= E[\max\{R_{N-1}, \hat{H}_{N-1}\}|\mathcal{H}_{N-1}] \\ &> \max\{R_{N-1}, E(\hat{H}_{N-1}|\mathcal{H}_{N-1})\} \\ &= \max\{R_{N-1}, E(E(V_N|\mathcal{H}_{N-1})|\mathcal{H}_{N-1})\} \\ &= V_{N-1}. \end{aligned}$$

In general, we believe that the expected value and variance of $\hat{V}_0$ has the forms shown in (5.5) and (5.6):

$$E(\hat{V}_0) = V_0 + a/m + \mathrm{O}(1/M^2), \tag{5.5}$$

$$\mathrm{Var}(\hat{V}_0) = \nu^2/M + \mathrm{O}(1/M^2). \tag{5.6}$$

In our application (Section 6), we give evidence that $\hat{V}_0$ behaves according to (5.5) and (5.6). Let $V(M_1)$ denote the estimator of V_0 based on M_1 replications. Consider, the following estimator:

$$\hat{V}_0^{\mathrm{unb}} = \frac{M_1 \times V(M_1) - M_2 \times V(M_2)}{M_1 - M_2}. \tag{5.7}$$

We find that $E(\hat{V}_0^{\mathrm{unb}}) = V_0 + \mathrm{O}(1/M^2)$ and so $\hat{V}_0^{\mathrm{unb}}$, is nearly unbiased. Let $\alpha = M_1/(M_1 + M_2)$ and suppose $V(M_1)$ is independent of $V(M_2)$, then

$$\mathrm{Var}\left(\hat{V}_0^{\mathrm{unb}}\right) = \frac{(\nu^2/M)}{(1-2\alpha)^2} + \mathrm{O}(1/M^2). \tag{5.8}$$

This variance indicates that a larger variance for the estimator is the penalty that we must pay for removing the first-order bias. But if we let $\alpha = 0.10$ then the standard deviation only increases by 25%, an acceptable amount.

5.2. Regression with q-splines

In this section, we describe how to calculate of $\hat{H}_n(\cdot)$ for $n = 1, \ldots, N-1$ using a q-spline when $\boldsymbol{H}_n = S_n$ is price of a stock at time n. Consult Seber and Wild (1989) for more details about regression splines. For a general treatise on splines, read Schumaker (1981). Let $S_{n,m}$ denote the observed stock price at time n from replication $m = 1, \ldots, M$. Let $\hat{H}_n(\cdot)$ denote an estimate of $H_n(\cdot)$, where $H_n(S_n) \equiv E(V_{n+1}|S_n) = E(V_{n+1}|\mathcal{H}_n)$. To start the algorithm, we let

$$\hat{V}_N(s) \equiv v_N \times I_N(s). \tag{5.9}$$

For $n = 1, \ldots, N-1$, we define

$$\hat{V}_{n+1}(s) = \max\{I_{n+1}(s), \hat{H}_{n+1}(s)\}. \tag{5.10}$$

Consider the simulated data $\{S_{n,m}, \hat{V}_{n+1}(S_{n+1,m})\}$. Let $\min_m\{S_{n,m}\} \equiv \kappa_0 < \kappa_1 < \cdots < \kappa_{D-1} < \kappa_D \equiv \max_m\{S_{n,m}\}$ denote the knots for the spline. In the next section, the knots were chosen so that the number of observations between two knots is nearly constant. The formula for the q-spline is

$$\hat{H}_n = \sum_{j=0}^{q} \phi_j \times s^j + \sum_{d=1}^{D-1} \xi_d \times \max\{0, (s-\kappa_d)^q\}. \tag{5.11}$$

This spline is a piecewise polynomial of degree q with $q-1$ continuous derivatives at the interior knots. The parameters $\phi_0, \ldots, \phi_q$ and $\xi_1, \ldots, \xi_{D-1}$ are those values that minimize the sum of squares

$$\begin{aligned} &\mathrm{SS}(\phi_0, \ldots, \phi_q, \xi_1, \ldots, \xi_{D-1}) \\ &= \sum_{m=1}^{M} \left(\hat{V}_{n+1}(S_{n+1,m}) - \hat{H}_n(S_{n,m})\right)^2. \end{aligned} \tag{5.12}$$

This estimation problem can easily be put in matrix form because (5.11) is linear in the parameters. Therefore, the solution that minimizes (5.12) can readily be found with classical linear regression formulas.

5.3. Regression with a local polynomial smoother

Another way to calculate of $\hat{H}_n(\cdot)$ is to use local polynomial regression. Consult Cleveland et al. (1988) for more details about this method. The advantage of this method over splines is that it requires fewer calculations, making it faster. In this case, we let

$$\hat{H}_n(s) = \sum_{j=0}^{q} \phi_j(s) \times s. \tag{5.13}$$

The parameters $\phi_0(s), \ldots, \phi_q(s)$ are those values that minimize the sum of squares

$$\begin{aligned} &\mathrm{SS}\ (\ \phi_0(s), \ldots, \phi_q(s)) \\ &\quad = \sum_{m\in\Gamma(s)} \left(\hat{V}_{n+1}(S_{n+1,m}) - \hat{H}_n(S_{n,m})\right)^2, \end{aligned} \tag{5.14}$$

where $\Gamma(s)$ is an index set equal to

$$\Gamma(s) = \{m = 1, \ldots, M : d(s) < S_{n,m} - s < e(s)\}, \tag{5.15}$$

and $0 < d(s) < e(s)$ are parameters that define the local region of regression. Estimation is straightforward because this problem fits the classical linear regression paradigm.

6. Approximating the value of an American put option

In this section, we illustrate out methods with an American put option. In this case $I_n = \max\{0, K - S_n\}$. We let $N = 12$, $K = 45$, $S_0 = 40$, $i = 0.07$, $\delta = \log_e(1+i)$, $v = \exp\{-\frac{1}{4}\delta n\}$, $\sigma = 0.3$, and

$$S_n = S_{n-1} \times \exp\{\tfrac{1}{4}\delta - \tfrac{1}{8}\sigma^2 + \tfrac{1}{2}\sigma Z_n\},\quad n = 1, \ldots, 12, \tag{6.1}$$

and $Z_n \sim N(0,1)$, and $Z_1, Z_2, \ldots$ are stochastically independent. In this example, n is the number of quarter years while δ is the annual force of interest. Moreover, we let $\boldsymbol{H}_n = S_n$ and $\mathcal{H}_n \equiv \sigma\{S_k \mid 0 \le k \le n\}$. Clearly, the process $\{S_n\}_{n=0}^{12}$ is Markovian. Moreover, the process $v_n \times S_n$ is a martingale because $E\{v_n \times S_n \mid \mathcal{H}_{n-1}\} = S_{n-1}\exp\{-\frac{1}{4}\delta(n-1)\}E\{\exp(-\frac{1}{8}\sigma^2 + \frac{1}{2}\sigma Z_n)\} = v_{n-1} \times S_{n-1}$.

Let $S_{n,m}$ denote the observed stock price at time n from replication $m = 1, \ldots, M$. For $n = 0$, we find that $V_0 > R_0$ and $V_0 = E(V_1 \mid \mathcal{H}_0) = E(V_1 \mid S_0) = H_0$. So, we let

$$\hat{V}_0 = \frac{1}{M}\sum_{m=1}^{M} \hat{V}_1(S_{1,m}). \tag{6.2}$$

The function $\hat{V}_1(s)$ is found recursively using (5.9) and (5.10). Consulting (5.3), the approximate optimal stopping time is

$$\hat{\tau}_0 = \sum_{k=0}^{N} k \times 1\left\{\hat{H}_h(S_h) > v_h I_h(S_h) \text{ for } 0 \le h < k \text{ and } \hat{H}_k(S_k) \le v_k I_k(S_k)\right\}. \tag{6.3}$$

Using (5.4) along with (6.3), evaluated at each of the sample paths, yields an estimator, $\tilde{V}_0$, that exhibits a smaller bias than $\hat{V}_0$ when M is large. This point is investigated in the next section.

6.1. Comparison of estimators

We estimated V_0 four different ways. Let $\hat{V}^l$ denote an approximation of V_0 based on Eq. (6.2) using a linear spline ($q = 1$, $D = 7$) and let $\hat{V}^q$ denote an approximation of V_0 based on Eq. (6.2) using a quadratic spline ($q = 2$, $D = 7$). Next, let $\tilde{V}^l$ denote an approximation of V_0 based on Eq. (5.4) using a linear spline ($q = 1$, $D = 7$) and let $\tilde{V}^q$ denote an approximation of V_0 based on Eq. (5.4) using a quadratic spline ($q = 2$, $D = 7$). We generated these estimates with $M = 4000$ sample paths of the process. To examine the properties of these estimators we repeated the experiment $\Gamma = 1000$ times and we generated the statistics $\hat{V}_\gamma^l$, $\hat{V}_\gamma^q$, $\tilde{V}_\gamma^l$, $\tilde{V}_\gamma^q$ for $\gamma = 1, \ldots, \Gamma$. The four statistics were calculated with the same simulated data. Let $\hat{V}_\gamma$ denote any of the four estimates. Then an estimate of $E(\hat{V}_\gamma)$ is

$$\bar{V} = \frac{1}{\Gamma}\sum_{\gamma=1}^{\Gamma} \hat{V}_\gamma, \tag{6.4}$$

while an estimate of the standard deviation $\sqrt{\mathrm{Var}(\hat{V}_\gamma)}$ is

$$\hat{v} = \sqrt{\frac{1}{\Gamma}\sum_{\gamma=1}^{\Gamma}(\hat{V}_\gamma - \bar{V})^2}. \tag{6.5}$$

Using (6.4) and (6.5), we calculated Z-statistics for the two-sided hypothesis that the four estimators are equal to the true value of $V_0 = 7.941$ (Tilley, 1993). The results are summarized in Table 1. All computations for this simulation

Table 1
Tests of biasedness

Null hypothesis	$\bar{V}$	$\hat{\nu}$	$\lvert Z \rvert$	Decision
$E(\hat{V}^l) = 7.941$	7.9507	0.0937	3.28	Reject
$E(\hat{V}^q) = 7.941$	7.9481	0.0933	2.40	Reject
$E(\tilde{V}^l) = 7.941$	7.9398	0.0991	0.38	Accept
$E(\tilde{V}^q) = 7.941$	7.9413	0.0998	0.10	Accept

were done with the computer language GAUSS. All decisions are based on a 5% level of significance. Examining Table 1, we can conclude that the $\hat{V}$ estimators are biased but we cannot reject the hypothesis of unbiasedness for the $\tilde{V}$ estimators. Actually $\tilde{V}$ is biased but when M is large the bias is very small. We did another experiment with $\Gamma = 2500$ and $M = 500$. In this case we found $\bar{V} = 8.030$ and $\hat{\nu} = 0.293$ and we rejected the null hypothesis $E(\tilde{V}^q) = 7.941$. It also seems that the quadratic estimators are different than the linear ones. Table 2 shows that this conclusion is supported by the data. The Z-statistic was calculated the same way as before except that the mean and standard deviation calculations were based on the differences, $\hat{V}^l_\gamma - \hat{V}^q_\gamma$, for example.

6.2. *Properties of the estimator* $\hat{V}^q$

Let us show that the estimator, $\hat{V}^q$, behaves according to (5.5) and (5.6). To examine the properties of this estimator we repeated the experiment $\Gamma = 7200$ times and we calculated $\bar{V}_M$ and $\hat{\nu}_M$ using (6.4) and (6.5) assuming that we had $M = 500, 1000, \ldots, 5000$ sample paths. These statistics are analyzed in Fig. 1. This graph reveals that $(\bar{V}_M - V_0)^{-1} \approx M/a$ and $(\hat{\nu}^2_M)^{-1} \approx M/\nu^2$, as hypothesized. This behavior was also observed for the estimator $\tilde{V}^q$. All graphs were done with the computer language GAUSS.

6.3. *Illustration of the quadratic spline*

Fig. 2 illustrates the shape of the quadratic spline with 12 graphs. Specifically, this figure plots each quadratic spline function, $\hat{H}_n(\cdot), n = 1, \ldots, 11$. Moreover, this figure also plots the paired observations $\{S_{n,m}, \hat{V}_{n+1}(S_{n+1,m})\}$ for $n = 1, \ldots, 11$ and $m = 1, \ldots, M = 1000$, that were used for estimating the spline. For $n = 12$ we plotted the pairs $\{S_{12,m}, v_{12} \times I_{12}(S_{12,m})\}$ for $m = 1, \ldots, M = 1000$.

6.4. *Unbiased estimation*

In this section, we investigate the unbiased estimator $\hat{V}_0^{\text{unb}}$, given in (5.7). In that formula, the biased estimator V denotes the estimator $\tilde{V}^q$. In our experiments, M_2 was always fixed at 500 while M_1 increased from 1000 to 5000 in increments of 500. To calculate $\bar{V}$ and $\hat{\nu}$, we repeated the experiment $\Gamma = 2500$ times. The results are given in Table 3. At the 5% level of significance, we found that we could never reject the two-sided hypothesis $E(\hat{V}_0^{\text{unb}}) = 7.941$. In conclusion, we find that (5.7) yields unbiased estimates.

6.5. *A comparison of the spline and local regression methods*

In conclusion, we compared the quadratic spline estimator $\tilde{V}^q$, based on formula (5.11), with the local quadratic estimator $\tilde{V}^{\text{local}}$, based on formula (5.13). We generated estimates with $M = 2000$ sample of the process. To

Table 2
Tests of equality of means

Null hypothesis	$\lvert Z \rvert$	Decision
$E(\hat{V}^q) = E(\hat{V}^l)$	9.25	Reject
$E(\tilde{V}^q) = E(\tilde{V}^l)$	2.49	Reject

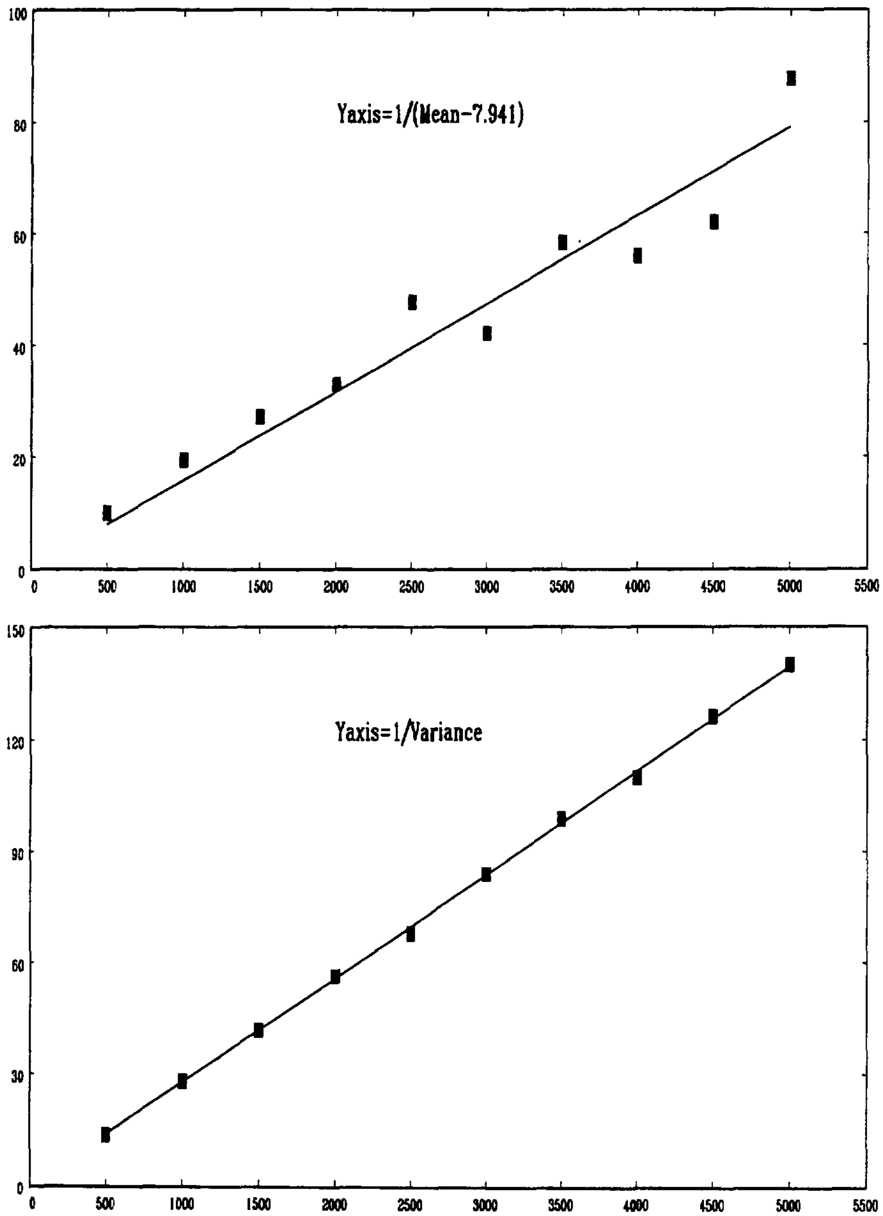

Fig. 1. A plot of $(\bar{V}_M - V_0)^{-1}$ (y-axis) and $(\hat{v}_M^2)^{-1}$ (y-axis) for $M = 500, \ldots, 5000$ (x-axis). Regression lines are also shown.

examine the properties of these estimators we repeated the experiment $\tau = 100$ times and we calculated $\tilde{V}_\gamma^q$ and $\tilde{V}_\gamma^{\text{local}}$ for $\gamma = 1, \ldots, \Gamma$. The two statistics were calculated with the same simulated data. To test the null hypothesis $E\{\tilde{V}^q\} = E\{\tilde{V}^{\text{local}}\}$, we calculated a Z-statistic that was based on the differences, $\tilde{V}_\gamma^q - \tilde{V}_\gamma^{\text{local}}$. We found that $|Z| = 0.88$ and concluded that the two methods were similar.

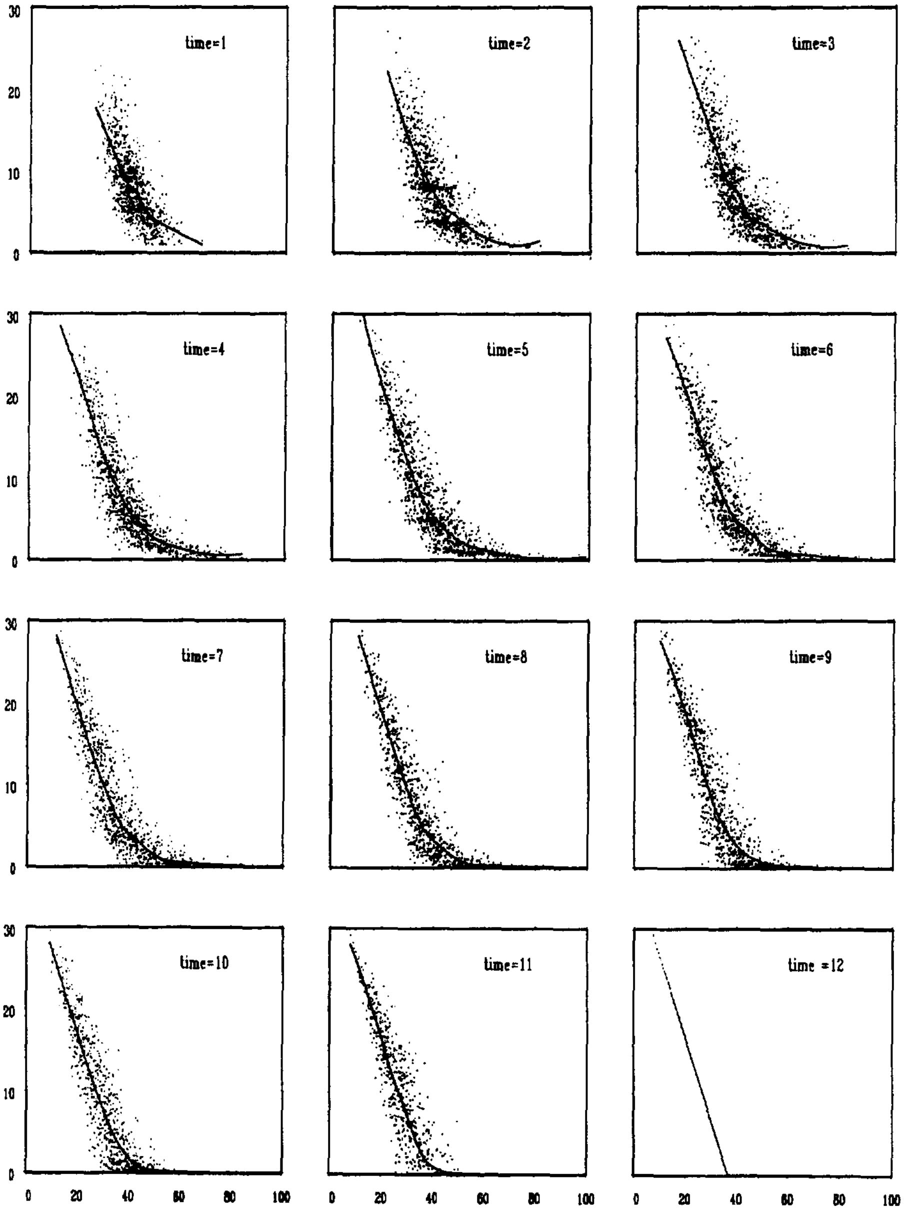

Fig. 2. A plot of the quadratic spline, $\hat{H}_n(\cdot)$, and the observations $\{S_{n,m}, \hat{V}_{n+1}(S_{n+1,m})\}$. The spline function is shown as a solid line while paired data is given as a scatterplot.

7. Summary

The article starts by defining the intrinsic value of various derivative securities like options. Next, the paper shows how to value cash flows when investors are risk neutral. This valuation method is consistent with the *Local Expectations Hypothesis* of the term structure of interest rates. Next, we give a backward induction algorithm for optimal stopping in finite discrete time and we relate it to the valuation of the early exercise privilege for American

Table 3
Tests of the hypothesis, $E(\hat{V}_0^{\mathrm{unb}}) = 7.941$

M_1	V	$\hat{\upsilon}$	$\|Z\|$	Decision
1000	7.9415	0.5110	0.04	Accept
1500	7.9482	0.2901	1.23	Accept
2000	7.9364	0.2137	1.08	Accept
2500	7.9382	0.1774	0.79	Accept
3000	7.9431	0.1532	0.69	Accept
3500	7.9416	0.1349	0.23	Accept
4000	7.9405	0.1240	0.19	Accept
4500	7.9397	0.1167	0.55	Accept
5000	7.9387	0.1074	1.09	Accept

options. Using this result, we show that the formula for an American call option is the same as the formula for the european call option, assuming no arbitrage opportunities exist.

Also, we present a simulation algorithm for approximating the value of the early exercise option for derivative securities. The algorithm uses a regression function to approximate conditional expectations. We also present an unbiased algorithm for approximating the value of the early exercise option. Finally, we illustrate the simulation and regression techniques with an American put option.

References

[1] Boyle, P.P. (1992). *Options and the Management of Financial Risk.* Society of Actuaries, Schaumburg, Il.

[2] Chow, Y.S., H. Robbins and D. Siegmund (1971). *Great Expectations: The Theory of Optimal stopping.* Houghton Mifflin, New York, NY.

[3] Cleveland, W.S. and S.J. Devlin (1988). Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association* 83 (403), 596–610.

[4] Cleveland, W.S., S.J. Devlin and E.Grosse (1988). Regression by local fitting-methods, properties and computational algorithms, *Journal of Econometrics* 37, 87–114.

[5] Cox, J.C., J.E. Ingersoll and S.A. Ross (1981). A re-examination of traditional hypothesis about the term structure of interest rates, *Journal of Finance* XXXVI, (4), 769–799.

[6] Härdle, W. (1980). *Applied Nonparametric Regression,* Cambridge University Press, New York, NY.

[7] Schachermayer, W. (1992). A Hilbert space proof of the fundamental theorem of asset pricing in finite discrete time, *Insurance: Mathematics and Economics* 11, 249–257.

[8] Schumaker, L.L. (1981). *Spline Functions Basic Theory.* Wiley, New York, NY.

[9] Seber, G.A.F. and C.J. Wild (1989). *Nonlinear Regression.* Wiley, New York, NY.

[10] Tang, K.S. and K.R. Vetzal (1994). Early exercise regions for exotic options, *Institute of insurance and Pension Research.* University of Waterloo, Ont.

[11] Tiley, J.A. (1993). Valuing American options in a path simulation model, *Transactions*, Vol. XLV. Society of Actuaries, Schaumburg, Il.