VALUATION OF THE RESET OPTIONS EMBEDDED IN SOME EQUITY-LINKED INSURANCE PRODUCTS

Phelim P. Boyle,* Adam W. Kolkiewicz,† and Ken Seng Tan††

ABSTRACT

This paper proposes a method for valuing American options using a Monte Carlo simulation approach. Our approach can be used to price the reset feature found in some equity-linked insurance contracts. We model this feature as a multiple shout option and give examples based on certain equity-linked insurance products that are very popular in Canada. These contracts are known as segregated fund contracts and the valuation of the embedded options in these contracts has posed serious challenges for actuaries. One of the advantages of the Monte Carlo approach in this connection is that it can be extended to handle different investment assumptions as well as multiple assets. We show how to modify the stochastic mesh model of Broadie and Glasserman (1997b) to incorporate quasi-Monte Carlo in the simulation and thus improve the efficiency. We benchmark the efficiency gains in our method using standard American options and multiple shout options.

1. INTRODUCTION

In the last two decades, insurance companies have introduced several new types of investment linked products. These products often include specific investment performance guarantees, and they have become very popular in the United States and Canada. In the United States, equity-indexed annuities provide benefits based directly on the performance of some broad-based equity portfolios such as Standard and Poor's (S&P) 500. The policyholders' benefit often takes the form of some percentage of the return on the reference portfolio. The financial guarantees found in these contracts can be viewed as embedded options, and a rich variety of such options is found in the marketplace.

In Canada, the corresponding contracts are known as segregated funds because the assets backing these contracts are segregated from the other assets of the insurance company. Initially, the segregated fund contracts were very similar to standard mutual funds except that they guaranteed that the policyholder would receive a minimum benefit after a period of 10 years. Typically this was expressed as a fixed percentage of the premium. Common levels of these percentages were 75% or 100%. This guarantee means that the insurance company has sold the policyholder a long-term put option.

In the early 1990s, Canadian insurers began issuing more complicated guarantees. Policyholders were given the option to reset the level of the guarantees periodically during the term of the contract. This is known as the rest option, and it corresponds to a shout option, which has been discussed in the finance literature (cf. Thomas 1993). Some Canadian companies also permitted the policyholder to switch from one portfolio to another several times during the duration of the contract. It is clear that many of these new features were added to give a marketing edge and that the issuing companies did not fully appreciate the financial implications of these additional features.

---

* Phelim P. Boyle, F.I.A., F.C.I.A., Ph.D., is Director of the Centre for Advanced Studies in Finance, University of Waterloo, Waterloo, ON N2L 3G1, Canada, e-mail: pboyle@uwaterloo.ca.

† Adam W. Kolkiewicz, Ph.D., is in the Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, ON N2L 3G1, Canada, e-mail: wakolkiewicz@uwaterloo.ca.

†† Ken Seng Tan, A.S.A., Ph.D., is in the Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, ON N2L 3G1, Canada, e-mail: kstan@uwaterloo.ca.

Because equity-linked products and segregated fund business use concepts from modern finance, their valuation and risk management have posed difficult challenges for actuaries. It is encouraging to note that the actuarial profession is redesigning the basic educational programs to include ideas from financial economics. We hope this paper contributes to this educational process because it uses and extends a finance-based model to value reset options.

Our approach to valuing these reset guarantees is to model them as multiple shout options and use a quasi-Monte Carlo simulation method to value them. To explain the method, we assume that the policyholder exercises these reset options optimally based on current market conditions. There are certainly other factors, economic and noneconomic, that may have an impact on the decision and should be taken into account when trying to identify the optimal strategy. However, these may depend on additional features of the segregated fund contract, like the possibility of an early withdrawal or the deferral of the maturity date, each time the guarantee is reset.

We now describe in general the nature of the contract that we will consider in this paper. Let $\{S_t, t \geq 0\}$ be the prices of the underlying fund and G be the amount guaranteed at maturity T. Then the value of the contract can be expressed as

$$\max\{G, S_T\} = S_T + \max\{G - S_T, 0\}.$$

The expression on the right-hand side implies that the value of the contract can be perceived as a sum of two components: the underlying fund and an option that pays the difference $G - S_T$ in the event the fund at time T falls below the guaranteed level. When G is a prespecified guarantee level, the option component becomes a standard put option, and its value can be computed using the celebrated Black-Scholes formula.

For the embedded options in the equity-linked insurance or segregated funds, the guaranteed level G is a complicated function that often depends on the history of the underlying fund. More specifically, the holder of the option is allowed to reset the guaranteed level up to m times during the life of the contract, where typically $m > 1$. If we denote by $\{\tau_1, \ldots, \tau_m\}$ the times when the option is reset, then G is calculated as the maximum of the initial guaranteed floor level K and the prevailing asset levels at the time of reset, that is,

$$G = \max\{K, S_{\tau_1}, \ldots, S_{\tau_m}\}.$$

Hence the value of the embedded option can be represented as $\max\{G - S_T, 0\}$. In the special case when $m = 1$, this corresponds to the payoff function of a shout option (Thomas 1993). For $m > 1$, we refer to this type of contract as the multiple shout option (see Cheuk and Vorst 1997).

To value this contract, we assume that the policyholder exercises these reset options optimally, and this means that we are treating the option as an American option. An American option is one that can be exercised at any time before its maturity. The valuation of the early exercise feature assumes that the optionholder acts to maximize the value of his or her contract. In the present context, it is sometimes argued that policyholders often act irrationally and that policyholder behavior is unpredictable. Nevertheless, we feel that a valuation based on an optimal exercise strategy is of interest.

First, from a risk management viewpoint strategy, it is prudent to know the full extent of the risk we have assumed. Second, in the age of the Internet and increasing consumer awareness, it is not hard to imagine an intermediary being set up to provide advice to consumers on the optimal exercise strategy (for a small fee). Third, we can only measure the extent of the so-called irrationality of a policyholder's decision if we have already computed the optimal strategy. Thus, our objective is to determine the maximal costs the issuer of the contract will incur because of the reset feature. By definition, any other reset strategy followed by the policyholders will result in smaller costs for the insurance company.

In addition, we should also take into account the computational errors arising from simulation. An interesting result established by Broadie and Glasserman (1997a) is that, if a simulation technique is used to find the price of an American option, the estimator is always biased. In view of this, the objective in this paper is to consider a simulation technique that produces an accurate estimate and will bound the true value from above.

The Monte Carlo method has powerful advantages in the present context. The main one is its flexibility. It can handle any investment assumption; it can be extended to deal with multiple sources of uncertainty and thus cope with a whole bundle of option-like features.[1] Other methods such as the lattice method and the partial differential equation approach (Windcliff, Forsyth, and Vetzal 1999) may be more efficient for subclasses of problems with a small number of state variables, but none can match the Monte Carlo method in its wide range of applicability. Until 1993, when Jim Tilley published his seminal paper showing that the Monte Carlo method could be used to price American options, the conventional wisdom was that the Monte Carlo technique could only be used to value European derivatives. Since then several papers have discussed the application of the Monte Carlo method to the valuation of American options. These include Barraquand and Martineau (1995); Broadie and Glasserman (1997a, b); Carrière (1996); Carr and Yang (1997, 1998); Grant, Vora, and Weeks (1997); Ibanez and Zapatero (1999); and Raymar and Zwecher (1997).

Despite this progress, the valuation of American options by simulation still remains a challenging technical problem and even the best methods developed to date are still not very efficient. The basic source of the difficulty is that the American option problem arises as a dynamic programming problem where we work backward, while the Monte Carlo method involves a forward-stepping evolution through time. From our perspective, the stochastic mesh method of Broadie and Glasserman (1997b) has the most desirable properties because it provides a confidence interval for the results. We modify their method in two ways. First, we make some technical changes to the stochastic mesh density function. Second, we show how a quasi-Monte Carlo method can be applied in this setting. The quasi-Monte Carlo method has been shown to have better computational properties than standard Monte Carlo for a range of European-style problems, especially for small- to medium-size problems.[2] We show that our modifications lead to considerable efficiency gains, even when the options contain early exercise features.

The layout of the rest of the paper is as follows. Section 2 provides a brief review of the pricing of American options by Monte Carlo simulation. In particular, we describe the stochastic mesh method of Broadie and Glasserman (1997b) as well as the enhancement method of Avramidis and Hyden (1999a, b). Section 3 describes the quasi-Monte Carlo method. The key idea here is to use carefully selected deterministic points rather than random points. These deterministic points have the property that they are very well distributed over the range of integration. Such sequences are known as low discrepancy sequences.

Section 4 explains our modifications to the stochastic mesh method for valuing American options. We present two sets of numerical examples in Section 5. First, we use our modified stochastic mesh approach to value standard American call options because we can benchmark our results with the accurate numbers. Second, we use our approach to value options with multiple shout opportunities and demonstrate that our approach leads to improved efficiency. The final section contains some brief conclusions and suggestions for future research.

2. PRICING AMERICAN OPTIONS USING SIMULATION

In this section, we review the pricing of American options using Monte Carlo simulation. In particular, we describe the stochastic mesh method of Broadie and Glasserman (1997b) because it forms the framework for our proposed valuation method. As noted earlier, the first paper to propose a solution to this problem was by Tilley (1993), who proposed an ingenious “bundling” algorithm to handle the early-exercise feature of the options. This was a path-breaking paper that stimulated additional research in the area. The original Tilley algorithm, however, does have sev-

---

[1] In the segregated funds application, we have to deal with both the reset feature and the switching option. This involves both multiple resets and the optimal fund selection problem. In addition, we may wish to model policyholder lapse behavior and use a range of stochastic investment assumptions, including stochastic interest rates.

[2] Broadie, Glasserman, and Ha (2000) have proposed a modification that is in the same spirit as ours but the details are different.

eral drawbacks. First, because of the nature of the algorithm, all the simulated paths must be stored and sorted at each exercise point. Because the simulation method typically requires a large number of paths to get a reasonably good estimate, this implies that the storage and sorting requirements can be significant. Second, there appears to be no obvious way to generalize the Tilley algorithm to higher dimensions. This is particularly unfortunate because the main advantage of using simulation is the ability to deal with many underlying state variables. Third, the Tilley algorithm is biased. It is not known how to estimate the bias or correct it.

Subsequently, significant progress has been made by Barraquand and Martineau (1995); Carrière (1996); Grant, Vora, and Weeks (1997); and Raymar and Zwecher (1997) in a more general setting. All these approaches however produce estimators that are biased—the same shortcoming as the Tilley algorithm. As computational effort increases, there is no guarantee that the result will converge to the true solution. See Boyle, Broadie and Glasserman (1997) for an example of where the Barraquand-Martineau method fails to converge. It is shown in Broadie and Glasserman (1997a) that, among a large class of simulation-based estimators, it is not possible to find an estimator of the American option that is unbiased. In other words, the estimators from simulation are always biased.

Recognizing this, Broadie and Glasserman (1997a, b) circumvent the problem by proposing to work with two estimators—one biased high and one biased low—and with the additional property that these two estimators are asymptotically unbiased. The true value must therefore necessarily be bounded by these two estimators. Furthermore the accuracy of the method can now be assessed from the confidence intervals constructed based on these two estimates.

The first algorithm proposed by Broadie and Glasserman (1997a) for computing the high-biased and low-biased estimators is the simulated tree method. The main criticism of this method is that the computational effort is of order nb^d, where n represents the number of underlying assets, b represents the number of branches at each node, and d is the number of exercise points. For practical implementation, d must not be too large or else the computational complexity grows exponentially. A typical segregated fund contract is very long-dated with many possible exercise dates. This makes the simulated random tree approach infeasible for this application.

In this paper, we consider the second algorithm proposed by Broadie and Glasserman (1997b). This technique is known as the stochastic mesh method. Similar to the simulated tree approach, this procedure has the desired property that it provides two estimators (biased high and biased low) and both are asymptotically unbiased. In addition, the computational effort for this algorithm only grows polynomially with the number of exercise points. This algorithm, therefore, provides a good starting point for tackling complex problems such as the valuation of options embedded in segregated funds.

The stochastic mesh algorithm consists of two components. The first involves generating a stochastic mesh for which the early-exercise boundary is approximated and consequently leads to a *mesh estimator,* which is a high-biased estimator. The second part of the algorithm requires generating independent stochastic paths of the underlying assets. Using the early-exercise boundary estimated from the stochastic mesh, a *path estimator,* which is low-biased, is derived. The following two subsections summarize the procedure for computing these estimators. A more complete description can be found in Broadie and Glasserman (1997b).

2.1 High-Biased Mesh Estimator

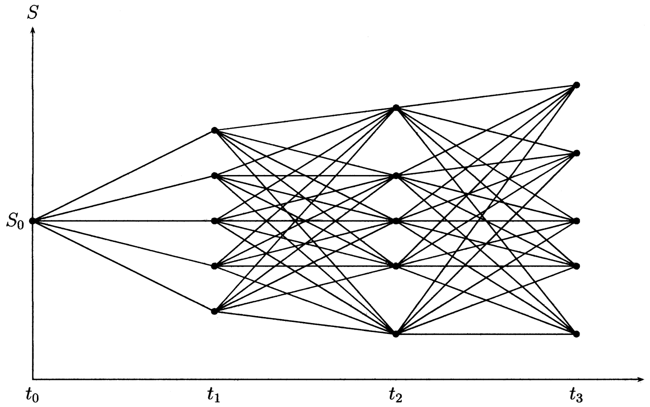
In this subsection, we describe the algorithm for obtaining the high-biased Broadie-Glasserman mesh estimator. To facilitate the discussion of the simulation technique, we assume that the option can only be exercised at $d + 1$ discrete time points $t_0 = 0, t_1 = 1, t_2 = 2, \ldots, t_d = T$. We also assume that the option depends on n underlying asset prices, with the initial asset prices denoted by the vector $\mathbf{S}_0 = (S_0^1, S_0^2, \ldots, S_0^n)$. Let $I(t, x) \geq 0$ denote the intrinsic value or the payoff from exercising the option at time t in state x and let $B(i, j)$ denote the discounting factor for the period (t_i, t_j); that is,

$$B(i, j) = e^{-r(t_j - t_i)},$$

where r is the annualized interest rate compounded continuously.

Figure 1

A Generic Mesh with $n = 1$, $b = 5$, $d = 3$



There are three steps in the computation of the high-biased mesh estimator:

Step 1. Starting at $\mathbf{X}_0(1) = \mathbf{S}_0$, generate $\mathbf{X}_t(i) = (X_t^1(i), X_t^2(i), \ldots, X_t^n(i))$ for $i = 1, 2, \ldots, b$; $t = 1, 2, \ldots, T$.

Step 2. At time $t_d = T$, set $\hat{Q}(T, \mathbf{X}_T(i))$ to be the terminal payoff of the option, assuming that it has not been exercised; that is,

$$\hat{Q}(T, \mathbf{X}_T(i)) = I(T, \mathbf{X}_T(i))$$

for $i = 1, 2, \ldots, b$.

Step 3. Recursively compute

$$\hat{Q}(t, \mathbf{X}_t(i)) = \max[I(t, \mathbf{X}_t(i)), \hat{C}(t, \mathbf{X}_t(i))] \quad (1)$$

for $i = 1, 2, \ldots, b$, and $t = T - 1, \ldots, 0$, where

$$\hat{C}(t, \mathbf{X}_t(i)) = \frac{B(t, t+1)}{b} \times \sum_{j=1}^{b} \hat{Q}(t+1, \mathbf{X}_{t+1}(j)) w(t, i, j) \quad (2)$$

and $w(t, i, j)$ is the weight attached to the arc joining $\mathbf{X}_t(i)$ to $\mathbf{X}_{t+1}(j)$.

Note that at time $t_0 = 0$ we have only one value of $i(=1)$ in Step 3. Hence $\hat{Q}(0, \mathbf{S}_0)$ is the mesh estimate for approximating the true value of the option. Figure 1 provides an illustration of the mesh for a one-dimensional case using $n = 1$, $b = 5$ and $d = 3$.

To implement the above algorithm, we need to specify how the random components $\mathbf{X}_t(\cdot)$ are generated and how the weights w on the arc are determined. Observe that $\hat{C}(t, \mathbf{X}_t(i))$ in (2) can be interpreted as the "holding value" or the "continuation value" from not exercising the option given that state $\mathbf{X}_t(i)$ is realized. Hence for arbitrary time t and state $\mathbf{S}_t = \mathbf{x}$, the continuation value can be expressed in terms of an expectation as follows:

$$C(t, \mathbf{x}) = B(t, t+1) \mathrm{E}[Q(t+1, \mathbf{S}_{t+1}) | \mathbf{S}_t = \mathbf{x}]. \quad (3)$$

As the stochastic mesh algorithm is carried out recursively, the expectation in the above equation must be estimated based on the information available from the mesh at time t with state $\boldsymbol{S}_t = X_t(i)$. In other words, $\hat{C}(t, \boldsymbol{x})$ must be estimated from $\{\hat{Q}(t+1, X_{t+1}(j)); j = 1, \ldots, b\}$. This can be accomplished by the following argument: Let $f(t, \boldsymbol{x}, \cdot)$ denote the density of $\boldsymbol{S}_{t+1}$ given $\boldsymbol{S}_t = \boldsymbol{x}$. Then we have

$$\begin{aligned} \mathrm{E}[Q(t+1, \boldsymbol{S}_{t+1}) | \boldsymbol{S}_t = \boldsymbol{x}] &= \int Q(t+1, \boldsymbol{u}) f(t, \boldsymbol{x}, \boldsymbol{u}) d\boldsymbol{u} \\ &= \int Q(t+1, \boldsymbol{u}) \frac{f(t, \boldsymbol{x}, \boldsymbol{u})}{g(t+1, \boldsymbol{u})} g(t+1, \boldsymbol{u}) d\boldsymbol{u} \\ &= \mathrm{E}\left[Q(t+1, \boldsymbol{U}) \frac{f(t, \boldsymbol{x}, \boldsymbol{U})}{g(t+1, \boldsymbol{U})}\right], \end{aligned} \quad (4)$$

where $\boldsymbol{U}$ is distributed according to a density $g(t+1, \cdot)$. This implies that, when the random components $X_t(i)$ for $i = 1, 2, \ldots, b$, and $t = 1, 2, \ldots, T$ are generated according to the density functions $g(t, \cdot)$, the expectation in Equation (4) can be approximated from $\hat{Q}(t+1, X_{t+1}(j))$ as

$$\mathrm{E}[Q(t+1, \boldsymbol{S}_{t+1}) | \boldsymbol{S}_t = \boldsymbol{x}] \approx \frac{1}{b} \sum_{j=1}^{b} \hat{Q}(t+1, X_{t+1}(j)) w(t, \boldsymbol{x}, X_{t+1}(j)), \quad (5)$$

where

$$w(t, \boldsymbol{x}, X_{t+1}(j)) = \frac{f(t, \boldsymbol{x}, X_{t+1}(j))}{g(t+1, X_{t+1}(j))}. \quad (6)$$

This provides a way of calculating $\hat{C}(t, \boldsymbol{x})$. The weight $w(t, \boldsymbol{x}, X_{t+1}(j))$ can be interpreted as the likelihood ratio or the Radon-Nikodym derivative.

In summary, the above results can be written more compactly as $\hat{Q}_h(t, X_t(i)) \equiv \hat{Q}(t, X_t(i))$ where

$$\hat{Q}_h(t, \boldsymbol{x}) = \max\left[I(t, \boldsymbol{x}), \frac{B(t, t+1)}{b} \times \sum_{j=1}^{b} \hat{Q}_h(t+1, X_{t+1}(j)) w(t, \boldsymbol{x}, X_{t+1}(j)) \right]. \quad (7)$$

This procedure is carried out recursively from $t = T, T-1, \ldots, 0$ so that the mesh estimator $\hat{Q}_h(0, \boldsymbol{S}_0)$ is obtained. The computational effort in generating the mesh is proportional to $n \times b \times d$ while the computational complexity in the recursive valuation (7) is proportional to $n \times b^2 \times d$. This implies that the overall computational complexity is polynomial in the state variables n, the mesh parameter b, and the number of exercise opportunities d.

Note that in Equation (7), we have explicitly introduced a subscript h to the notation Q to emphasize that the mesh estimator is biased high, as shown in Broadie and Glasserman (1997b, Theorem 1). This is a consequence of Jensen's inequality and can be explained as follows. Suppose that $\mathrm{E}[\hat{Q}(t+1, \boldsymbol{x})] \geq Q(t+1, \boldsymbol{x})$ for all $\boldsymbol{x}$. From Equation (2) and Jensen's inequality we have

$$\mathrm{E}[\hat{Q}(t, \boldsymbol{x})] \geq \max\{I(t, \boldsymbol{x}), \mathrm{E}[\hat{C}(t, \boldsymbol{x})]\}.$$

Because at each time step the mesh points are identically distributed, we have

$$\begin{aligned}
\mathrm{E}[\hat{C}(t, \boldsymbol{x})] &= B(t, t+1)\mathrm{E}[\hat{Q}(t+1, X_{t+1}(1)) w(t, \boldsymbol{x}, X_{t+1}(1))] \\
&= B(t, t+1)\mathrm{E}[\mathrm{E}[\hat{Q}(t+1, X_{t+1}(1)) | X_{t+1}(1)] \\
&\quad \times w(t, \boldsymbol{x}, X_{t+1}(1))] \\
&\geq B(t, t+1)\mathrm{E}[Q(t+1, \boldsymbol{S}_{t+1}) | \boldsymbol{S}_t = \boldsymbol{x}],
\end{aligned}$$

where the last inequality is a consequence of the initial assumption and Equation (4). Therefore,

$$\begin{aligned}
\mathrm{E}[\hat{Q}(t, \boldsymbol{x})] &\geq \max\{I(t, \boldsymbol{x}), B(t, t+1) \\
&\quad \times \mathrm{E}[Q(t+1, \boldsymbol{S}_{t+1}) | \boldsymbol{S}_t = \boldsymbol{x}\} = Q(t, \boldsymbol{x}),
\end{aligned}$$

and by mathematical induction, it follows that the mesh estimator $\hat{Q}_h(0, \boldsymbol{S}_0)$ is biased high. Furthermore, it can be shown that, under some moment conditions, the estimator is asymptotically unbiased as $b \to \infty$ (see Broadie and Glasserman 1997b, Theorem 2). The latter is an important result that ensures that, as computational effort increases (by increasing the mesh parameter b), the mesh estimator converges to the true value. This is in contrast to other algorithms such as Tilley (1993) and Barraquand and Martineau (1995) where the convergence cannot be guaranteed.

We now discuss the mesh density function $g(t, \cdot)$. Recall that the mesh points are generated according to this density function. Broadie and Glasserman noted that the choices for these functions "*are crucial to the success of the method*" (1997b, p. 5). A natural choice of the mesh density function is to set $g(t, \cdot) = f(t, \cdot)$ for $t = 1, 2, \ldots, T$, where $f(t, \cdot)$ is the marginal density of $\boldsymbol{S}_t$ given the initial value $\boldsymbol{S}_0$. This candidate, however, leads to estimators whose variance grows exponentially with T (or d) (Broadie and Glasserman 1997b, prop. 1). The choice proposed is to set

$$g^{BG}(t, \boldsymbol{u}) = \begin{cases} f(0, \boldsymbol{S}_0, \boldsymbol{u}), & \text{for } t = 1 \\ \dfrac{1}{b} \displaystyle\sum_{j=1}^{b} f(t-1, X_{t-1}(j), \boldsymbol{u}), & \text{for } t = 2, \ldots, T. \end{cases} \quad (8)$$

Note that $g^{BG}(t, \boldsymbol{u})$, $t = 2, \ldots, T$ is a mixture density. Broadie and Glasserman refer to the choice in Equation (8) as the average density function. The efficiency of the mesh estimator based on this average density function will be investigated in Section 5.

To conclude this subsection, we point out that the stochastic mesh method can also be applied to value European options. Although in this case there is no need to estimate the optimal exercise strategy, it is still necessary to proceed through the recursive valuation as follows:

$$\hat{Q}(t, X_t(i)) = \frac{B(t, t+1)}{b} \times \sum_{j=1}^{b} \hat{Q}(t+1, X_{t+1}(j)) w(t, \boldsymbol{x}, X_{t+1}(j)) \quad (9)$$

for $i = 1, 2, \ldots, b$, and $t = T-1, \ldots, 0$.

2.2 Low-Biased Estimator

We now describe the low-biased estimators. There are many ways of obtaining such estimators. The approach suggested by Broadie and Glasserman (1997b) is to exploit the early-exercise boundary implied from the mesh. First, an independent trajectory is simulated and the optimal stopping time along this trajectory is determined based on the early-exercise policy approximated from the mesh. Consequently, the payoff at the optimal stopping time is calculated and this procedure is repeated for many independent replications with the low-biased estimator defined as the average over these values. Broadie and Glasserman refer to this estimator as the path estimator.

More precisely, the procedure for calculating the path estimator can be described as follows:

Step 1. Simulate the trajectory $\boldsymbol{S} = (\boldsymbol{S}_0, \boldsymbol{S}_1, \ldots, \boldsymbol{S}_T)$ using the transition density function $f(t, \boldsymbol{x}, \cdot)$.

Step 2. Approximate the optimal exercising strategy using

$$\hat{\tau}(\boldsymbol{S}) = \min\{t : I(t, \boldsymbol{S}_t) \geq \hat{Q}(t, \boldsymbol{S}_t)\}, \qquad (10)$$

where $\hat{Q}(t, \boldsymbol{S}_t)$ is determined from the mesh using Equation (7).

Step 3. Record $\hat{q} = B(0, \hat{\tau}) I(\hat{\tau}, \boldsymbol{S}_{\hat{\tau}})$ and repeat Steps 1–3 until N_p independent trajectories have been simulated.

Step 4. The path estimator for the American option is given by averaging over these N_p estimates.

Because the stopping time approximated from Equation (10) is not necessarily an optimal stopping time, this implies that the path estimator is biased low (see Broadie and Glasserman 1997b, Theorem 3). Furthermore, in Theorem 4 Broadie and Glasserman also show that the path estimators estimated from the above algorithm are asymptotically unbiased. In terms of the computational effort for the path estimator, it is proportional to $n \times b \times d \times N_p$.

Avramidis and Hyden (1999a, b) suggest an alternate approach for obtaining a biased-low estimator by partitioning the mesh into two disjoint sets. These sets are then used recursively to estimate the optimal exercise policies and the continuation values. More precisely, let $\mathcal{X} = \{1, 2, \ldots, b\}$ be the set of indices of mesh points at each time period, and let $\mathcal{Y} \in \mathcal{X}$ denote an arbitrary subset of all indices, and $\mathcal{Y}'$ be its complement with respect to $\mathcal{X}$. Furthermore, let's define $\hat{C}(t, \boldsymbol{X}_t(i), \mathcal{Y})$ as the estimate of the continuation values at time t based only on the states in the subset $\mathcal{Y}$ from time $t + 1$; that is,

$$\hat{C}(t, \boldsymbol{X}_t(i), \mathcal{Y}) = \frac{B(t, t+1)}{|\mathcal{Y}|} \times \sum_{j \in \mathcal{Y}} \hat{Q}_l(t+1, \boldsymbol{X}_{t+1}(j)) w(t, \boldsymbol{X}_t(i), \boldsymbol{X}_{t+1}(j)), \qquad (11)$$

where w is defined in Equation (6) and $|\mathcal{Y}|$ is the number of elements in $\mathcal{Y}$. Note that, in contrast to Equation (2), the continuation value defined above depends explicitly on the mesh points in the set $\mathcal{Y}$. The estimate of the option $\hat{Q}_l(t, \boldsymbol{X}_t(i))$ is defined recursively as follows:

Step 1. At time $t_d = T$, set

$$\hat{Q}_l(T, \boldsymbol{X}_T(i)) = I(T, \boldsymbol{X}_T(i))$$

for $i = 1, \ldots, b$.

Step 2. Recursively compute

$$\hat{Q}_l(t, \boldsymbol{X}_t(i)) = \frac{1}{b} \sum_{j=1}^{b} \hat{Q}_l(t, \boldsymbol{X}_t(i), \mathcal{X}_{-j})$$

for $i = 1, 2, \ldots, b$ and $t = T - 1, \ldots, 0$, where

$$\hat{Q}_l(t, \boldsymbol{X}_t(i), \mathcal{X}_{-j}) = \begin{cases} I(t, \boldsymbol{X}_t(i)), & \text{if } I(t, \boldsymbol{X}_t(i)) \geq \hat{C}(t, \boldsymbol{X}_t(i), \mathcal{X}_{-j}) \\ \hat{C}(t, \boldsymbol{X}_t(i), \mathcal{X}'_{-j}), & \text{otherwise,} \end{cases} \qquad (12)$$

and $\mathcal{X}_{-j}$ is the set of all indices in $\mathcal{X}$ excluding the j-th element.

It is shown in Avramidis and Hyden (1999b) that the estimator $\hat{Q}_l$ is biased low, hence explaining the subscript l to the notation Q above.

The principal advantage of the above algorithm is that the low-biased estimator can be obtained concurrently with the high-biased estimator. This is in contrast to the biased low Broadie-Glasserman path estimator, which requires generating an independent set of trajectories, thus increasing the computational effort. This feature is exploited by Avramidis and Hyden (1999a) for proposing a so-called *average mesh estimator,* $\hat{Q}_\alpha$. Instead of recursively computing biased high and biased low estimators separately, they suggest computing a revised estimator by averaging the biased high and biased low estimators at each backward induction step. The revised algorithm can now be described as follows:

$$\hat{Q}_a(T, \boldsymbol{X}_T(i)) = I(T, \boldsymbol{X}_T(i))$$

for $i = 1, \ldots, b$; and for $t = T - 1, \ldots, 0$ and $i = 1, 2, \ldots, b$, the average mesh estimator is computed recursively as

$$\hat{Q}_a(t, \boldsymbol{X}_t(i)) = \tfrac{1}{2}[\hat{Q}_l(t, \boldsymbol{X}_t(i) + \hat{Q}_h(t, \boldsymbol{X}_t(i))]$$

where the values $\hat{Q}_\alpha(t + 1, \cdot)$ are used as substitutes for the values $\hat{Q}_l(t + 1, \cdot)$ and $\hat{Q}_h(t + 1, \cdot)$ in the calculation of $\hat{Q}_l(t, \cdot)$ and $\hat{Q}_h(t, \cdot)$, respectively.

It is intuitively clear in the case where the high and low biases are of the same magnitude, that using the average estimator could lead to a further reduction in the bias. The downside is that the bias of the resulting estimator is indeterminate. If we compute the biased high and biased low estimators separately, the $100(1 - \alpha)\%$ confidence interval can be constructed from N independent mesh estimators $\hat{Q}_{l,k}(0, \boldsymbol{S}_0)$ and $\hat{Q}_{h,k}(0, \boldsymbol{S}_0)$, $k = 1, \ldots, N$ as follows:

$$\left[\bar{Q}_l - z_{\alpha/2}\frac{\hat{\sigma}(\hat{Q}_l)}{\sqrt{N}}, \quad \bar{Q}_h + z_{\alpha/2}\frac{\hat{\sigma}(\hat{Q}_h)}{\sqrt{N}}\right], \quad (13)$$

where $z_{\alpha/2}$ is the $1 - \alpha/2$ quantile of the standard normal distribution and

$$\bar{Q}_l = \frac{1}{N}\sum_{k=1}^{N} \hat{Q}_{l,k},$$

$$\bar{Q}_k = \frac{1}{N}\sum_{k=1}^{N} \hat{Q}_{h,k},$$

and $\hat{\sigma}(\hat{Q}_l)$ and $\hat{\sigma}(\hat{Q}_h)$ are the sample standard deviations of $\hat{Q}_l$ and $\hat{Q}_h$, respectively. Observe that the confidence interval in Equation (13) is obtained by combining the lower confidence limit from the biased low estimator with the upper confidence limit from the biased high mesh estimator. This implies that the resulting interval is a valid $100(1 - \alpha)$ confidence interval for the true value Q because, for finite b, $\hat{Q}_l$ is biased low while $\hat{Q}_h$ is biased high. In fact, the confidence interval in Equation (13) will be conservative. In our numerical comparison, the above confidence interval is not constructed because we are primarily interested in assessing the bias arising from our proposed high-biased estimators and comparing that to the average mesh estimator of Avramidis and Hyden (1999a).

3. QUASI-MONTE CARLO METHODS AND LOW DISCREPANCY SEQUENCES

The principal criticism of the Monte Carlo method is its slow convergence. It achieves a convergence rate of $\mathcal{O}(N^{-1/2})$, which is independent of the dimension of the underlying problem. Several different methods for speeding up the convergence have been proposed. These techniques are known as variance reduction techniques. The quasi-Monte Carlo methods or low discrepancy (LD) methods were first introduced to the problem of derivative pricing by a number of authors including Joy, Boyle, and Tan (1996) and Paskov and Traub (1995). Since then, this technique has spurred considerable interest, as evidenced by the proliferation of subsequent publications: Acworth, Broadie, and Glasserman (1998); Berman (1997–98); Boyle and Tan (1997); Caflisch, Morokoff, and Owen (1997); Fishman, Fitton, and Galperin (1997); Galanti and Jung (1997); Mitchell (1998); Ninomiya and Tezuka (1996); Owen and Tavella (1997); Paskov (1996); Tan (1998); and Tan and Boyle (1997, 2000).

The LD methods rely on the use of specially selected deterministic sequences instead of random sequences. These deterministic sequences have the property that they are well dispersed throughout the required space and are known as low discrepancy sequences. The monograph by Niederreiter (1992) provides an excellent discussion of these sequences. It should be pointed out that, in practice, random sequences are also "deterministic" in that they are generated from deterministic algorithms, typically based on the linear congruential method.

The sequence of numbers from this generation mimics a sample of independent and identically distributed $U(0, 1)$ random variables. In this aspect, random sequences are commonly known as pseudo-random sequences and random number generators as pseudo-random number generators. The low discrepancy sequences, on the other hand, are generated so that they have greater uniformity at the expense of independence and randomness.

The idea of using well-dispersed sequences rather than randomly generated numbers is motivated by the classical result in one dimension. Suppose we are interested in evaluating the following integral:

$$\int_0^1 f(x)dx = \theta. \tag{14}$$

To solve this problem[3] using the Monte Carlo integration technique, a point is randomly drawn from [0, 1). The value evaluated at this sampled point is recorded. This procedure is repeated for N iterations so that the unbiased Monte Carlo estimator of θ is given by

$$\hat{\theta} = \frac{1}{N} \sum_{i=1}^{N} f(x_i),$$

where x_i is the i-th randomly selected point from the interval [0, 1). In practice, the points x_i, $i = 1, 2, \ldots, N$ are generated from a pseudo-random generator.

From the law of large numbers, the estimator $\hat{\theta}$ is unbiased and converges to the true value θ asymptotically. The central limit theorem also implies that the accuracy of the estimator $\hat{\theta}$ can be gauged by constructing confidence intervals. As mentioned earlier, an important property of the Monte Carlo method is that the rate of convergence is $\mathbb{O}(N^{-1/2})$, regardless of the dimensionality of the problem.

For reasonably smooth one-dimensional functions f, it is well known that a convergence rate much better than $\mathbb{O}(N^{-1/2})$ is possible by carefully choosing the points where the function is evaluated. These methods include the midpoint rule, Simpson's rule, and trapezoidal rule. For instance, in the case of the midpoint rule, the points are chosen according to the following scheme:

$$x_i = \frac{2i-1}{2N} \quad \text{for } i = 1, 2, \ldots, N.$$

This can achieve an order of convergence $\mathbb{O}(N^{-1})$, which is considerably better than the rate attained by the Monte Carlo method. Compared to the random points, the points from the midpoint rule are deterministic, with the property that they are equally spaced and *uniformly distributed* in the interval [0, 1). Such deterministic points can achieve a better rate of convergence than the randomly generated points. This has led researchers to believe that the uniformity criterion, rather than randomness property of the points, plays a greater role in determining the efficiency. An attempt has been made to generalize this result to higher dimensions. A notion of uniformity of the points in high dimensions has been developed. Deterministic sequences of s-dimensional points that have good "uniformity" in $[0, 1)^s$ are termed low discrepancy sequences. Using these sequences, one can achieve a convergence rate of $\mathbb{O}(N^{-1} \log^s N)$, which is superior to the Monte Carlo rate $\mathbb{O}(N^{-1/2})$ for large N. Conventionally, methods involving the low discrepancy sequences are misleadingly referred to as quasi-Monte Carlo methods, although there is nothing random about these sequences.

Explicit constructions of low discrepancy sequences are given by Faure (1982), Halton (1960), Niederreiter (1987), Sobol' (1967), and Tezuka (1993). Computer implementation of many of these sequences can be found in Bratley and Fox (1988); Bratley, Fox, and Niederreiter (1992); Fox (1986) and Shukhman (1994). These sequences have been applied to the computation of complex derivatives, and the general consensus is that they perform substantially better than Monte Carlo methods, even in very high dimensions such as 360. However, as pointed out in Tan (1998) and Tan and Boyle (2000), naive application can be dangerous and lead to erroneous estimates that are significantly worse than those produced by the Monte Carlo method. For instance, when applying LD methods to a problem that contains a discontinuity, the superior low discrepancy rate of convergence cannot be real-

---

[3] Of course, for many integrable functions f, the analytic value of the integral can be determined directly using elementary integration rules.

ized. Ways to rectify some of these problems are proposed in these papers.

In the next section, we will show how to unleash the power of LD methods to improve the Broadie and Glasserman method for valuing American options using simulations. To the best of our knowledge, this is one of the first applications of LD methods to the valuation of American options.

4. LOW DISCREPANCY MESH METHOD

In this section, we show how to improve the Broadie-Glasserman high-biased mesh estimator. This, in turn, depends on how we construct the mesh. As noted by Broadie and Glasserman (1997b), the choice for the mesh density functions $g(t, \cdot)$ plays a crucial role in the underlying efficiency of the method. For a given function $g(t, \cdot)$, the precision of the mesh estimator depends on how uniformly the random components $X_t(i)$, $i = 1, 2, \ldots, b$ span the space at each time t.

One candidate for the mesh density function suggested by Broadie and Glasserman (1997b) is the average density functions $g^{BG}(t, \cdot)$, defined in Equation (8). The justification for using $g^{BG}(t, \cdot)$ stems from the practicality of computing the European counterpart. In this case, the European option can be calculated directly (see Broadie and Glasserman 1997b, prop. 2) as

$$\frac{B(0, T)}{b} \sum_{i=1}^{b} I(T, X_T(i)), \tag{15}$$

which avoids the recursive valuation algorithm described in Equation (9). While the mesh density function $g^{BG}(t, \cdot)$ has computational advantages for calculating European options, we are primarily interested in the valuation of the American options. In any case, a recursive approach will always be required for estimating the value of American options. Hence, this computational benefit does not apply to the pricing of American-style options.

Another inherent difficulty with $g^{BG}(t, \cdot)$ relates to generating random samples from this distribution. According to step 1 of the algorithm in Section 2.1, the random vectors $X_t(\cdot)$ are distributed according to the mesh density functions $g(t, \cdot)$. When $g^{BG}(t, \cdot)$ is selected for the mesh density function, Broadie and Glasserman (1997b) outlined the following generation procedure:

For $t = 1$, the random components $X_1(i)$, $i = 1, \ldots, b$ are easily generated from $g^{BG}(1, \cdot) = f(0, S_0, \cdot)$, because $X_0 = S_0$. For $t > 2$, the situation is more complicated because $g^{BG}(t, \cdot)$ is a mixture of the conditional density functions $f(t-1, X_{t-1}(\cdot), \cdot)$. The random samples can be generated as follows:

- Initialize $i = 0$.
- Repeat
  1. Set $i = i + 1$.
  2. Randomly pick an integer i^* that is uniformly distributed in the set $\{1, 2, \ldots, b\}$.
  3. Generate $X_t(i)$ from $X_{t-1}(i^*)$, using the transition density $f(t-1, X_{t-1}(i^*), \cdot)$, until $i = b$.

It can be shown that the resulting random components $X_t(i)$, $i = 1, \ldots, b$, are distributed according to the average density function $g^{BG}(t, \cdot)$. In fact, the above procedure is a standard technique for generating random samples from mixture densities (see Devroye 1986). Hence, the random samples converge to $g^{BG}(t, \cdot)$ only at the Monte Carlo rate of $\mathbb{O}(N^{-1/2})$.

If one naively applies the low discrepancy points to the above procedure, one finds that the superior LD convergence rate is not attained. In some situations, the low discrepancy points could perform worse than random points. One might be tempted to conclude that LD sequences offer no computational advantages in the case of pricing American options. This observation may also explain why none of the LD papers to date (apart from the work of Broadie, Glasserman, and Ha 2000) has explored the feasibility of the methods to the pricing of American-style options. Recall that to generate a random normal variate, the Box-Muller transformation method (Press et al. 1992, p. 289) is often used. As cautioned by Joy, Boyle and Tan (1996, footnote 6), a direct inverse transformation must be used with the LD methods. Otherwise, the low discrepancy of the points will be distorted and the uniformity property will not be preserved. Hence, the computational efficiencies of the LD sequences are not attained. The same reasoning applies when one blindly employs LD points and uses the above procedure to generate samples from mixture densities.

Observe that, theoretically, in Equation (4) any density function $g(t, \cdot)$ can be used as long as

$g(t, \cdot) > 0$. To harness the power of the LD methods, the mesh density function must be selected judiciously. Our proposed choice of this density can be described as follows: At each time point t, the mesh density function $g(t, \cdot)$ has the same functional form as the transition density functions $f(t, \mathbf{x}, \cdot)$ except that the first two moments of $g(t, \cdot)$ match exactly the corresponding moments of the average density function $g^{BG}(t, \cdot)$. We denote this mesh function as $\tilde{g}(t, \cdot)$.

In the Black-Scholes framework with one asset, the transition density $f(t-1, X_{t-1}(i), \cdot)$ is lognormal and corresponds to a random variable of the form

$$X_{t-1}(i)e^{\mu+\sigma Z},$$

where Z is normally distributed with the mean and variance equal to zero and one. Then, it is easy to show that the first two moments of the average density $g^{BG}(t, \cdot)$ are equal to

$$M_1 = e^{(1/2)\sigma^2} \cdot \frac{1}{b} \sum_{i=1}^{b} e^{\zeta_i},$$

and

$$M_2 = e^{2\sigma^2} \cdot \frac{1}{b} \sum_{i=1}^{b} e^{2\zeta_i},$$

respectively, where $\zeta_i = \ln(X_{t-1}(i)) + \mu$, $i = 1, \ldots, b$. In this case, the mesh density $\tilde{g}(t, \cdot)$ is lognormal, and it corresponds to a random variable of the form

$$e^{\mu^* + \sigma^* Z},$$

where

$$\mu^* = \ln\left(\frac{M_1^2}{\sqrt{M_2}}\right)$$

and

$$\sigma^* = \sqrt{\ln\left(\frac{M_2}{M_1^2}\right)}.$$

Observe also that for $t = 1$, the proposed density is equal to $g^{BG}(t, \cdot)$ and is the same as the transition density.

There are at least three advantages of using $\tilde{g}(t, \cdot)$. First, it is very simple and can easily be implemented. Second, it takes into account the actual realization of the mesh point in the previous time period. Third, and most important, the superior convergence rate of the LD methods is preserved. When random sequences are applied to both $\tilde{g}(t, \cdot)$ and $g^{BG}(t, \cdot)$, the Monte Carlo convergence rates remain unaffected. However, when LD sequences are applied to the newly proposed $\tilde{g}(t, \cdot)$, the low discrepancy of the sequences will not be disrupted and the meshes X_t, $t = 1, 2, \ldots, T$ become more uniformly distributed. Consequently, this increases the precision of the mesh estimator. The enhanced precision is clearly demonstrated in the numerical examples presented in the following section.

5. NUMERICAL EXAMPLES

In this section, we consider several numerical examples to illustrate the efficiency of the proposed methods. To facilitate the evaluation, we use the simple Black and Scholes (1973) assumptions so our simulation results can be benchmarked with accurate values obtained from other methods.

5.1 American Call Options

In this subsection, we consider the valuation of American call options. Assume that the options depend only on a single asset with the following parameter values: strike price \$100, volatility 30%, risk-free rate 5%, dividend rate 10%, time until maturity three years, and three possible levels of initial stock prices—\$90, \$100 and \$110.[4] Further assume that there are either 10 or 50 evenly spaced discrete exercise points for each option. Such options are known as Bermudan option contracts because they are intermediate between European and American options.

For each option contract, we consider three different ways of creating the mesh. The first mesh is generated using $g^{BG}(t, \cdot)$, that is, the mixture density proposed by Broadie and Glasserman (1997b). From the resulting mesh, two option estimators are computed. The first is the high biased Broadie-Glasserman (BG) mesh estimator, $\hat{Q}_h(0, S_0)$. The second is the average mesh estimator, $\hat{Q}_a(0, S_0)$, considered by Avra-

---

[4] Note that we have considered an asset that pays dividends; hence, early exercise might be optimal for the American call options.

midis and Hyden (1999a). We label this estimator as the AH mesh estimator.

The second type of mesh is generated using the proposed $\tilde{g}(t, \cdot)$ together with the LD methods. For the low discrepancy sequences, we used the randomized Faure sequences proposed by Owen (1996). See also Tan and Boyle (2000) for application of these sequences for pricing high-dimensional derivative securities. We denote the resulting biased high estimate of the option as the LD mesh estimator. The last type of mesh, which is only of expository interest, also uses the proposed density $\tilde{g}(t, \cdot)$ to generate the mesh points. The only difference is that in this case random sequences, instead of low discrepancy sequences, are used. We label these estimators as MC mesh estimators.

To provide a better assessment of how the various estimators converge and also to evaluate the bias of these estimators, we construct the mesh independently 10 times to produce 10 option estimators for each method and for each mesh size $b \in \{256, 1024, 4096\}$. For the mesh based on the Monte Carlo methods, we use the "ran2" pseudorandom number generator of Press et al. (1992). For the mesh based on LD methods, independent randomized low discrepancy sequences are produced from randomly generated independent sets of permutation functions. These sequences, in turn, are used to simulate independent meshes. This also explains why we use the randomized sequences in this study and not the classical sequences such as Faure (1982), Halton (1960), Niederreiter (1987), Sobol' (1967) and Tezuka (1993).

The probabilistic error bounds would not have been possible if we had used the traditional low discrepancy sequences. The resulting sample mean (as well as its sample standard errors) based on the 10 independent option estimates are tabulated. The results are summarized in Tables 1 and 2. In addition to reporting the simulation results for each method, the American call option estimates from the binomial lattice are also included. These values are used as a benchmark for comparison purposes. Our conclusions, based on the simulation results, are as follows:

- First the BG mesh estimators converge monotonically from above. This is consistent with the

Table 1

Various Mesh Estimators for American Call Options with 10 Exercise Points

			Mesh Estimators			
S	σ	b	BG	AH	LD	MC
90	20%	256	5.18 (0.29)	4.79 (0.29)	4.41 (0.01)	5.12 (0.23)
		1,024	4.78 (0.14)	4.47 (0.14)	4.39 (0.00)	4.66 (0.19)
		4,096	4.50 (0.07)	4.44 (0.05)	4.39 (0.00)	4.51 (0.07)
		Lattice	4.39	4.39	4.39	4.39
90	40%	256	17.26 (1.04)	15.76 (1.04)	14.31 (0.02)	16.97 (0.71)
		1,024	15.92 (0.50)	14.81 (0.52)	14.24 (0.00)	15.56 (0.68)
		4,096	14.77 (0.25)	14.57 (0.23)	14.24 (0.00)	14.89 (0.28)
		Lattice	14.23	14.23	14.23	14.23
100	20%	256	8.97 (0.33)	8.42 (0.34)	8.00 (0.01)	8.94 (0.30)
		1,024	8.48 (0.17)	8.21 (0.16)	7.98 (0.00)	8.24 (0.21)
		4,096	8.07 (0.06)	8.00 (0.05)	7.99 (0.00)	8.02 (0.07)
		Lattice	7.98	7.98	7.98	7.98
100	40%	256	22.59 (1.16)	20.96 (1.13)	19.09 (0.02)	22.35 (0.83)
		1,024	21.00 (0.59)	19.89 (0.56)	19.02 (0.00)	20.48 (0.76)
		4,096	19.54 (0.26)	19.30 (0.24)	19.02 (0.00)	19.62 (0.30)
		Lattice	19.02	19.02	19.02	19.02
110	20%	256	14.26 (0.33)	13.70 (0.33)	13.19 (0.01)	14.03 (0.31)
		1,024	13.67 (0.14)	13.42 (0.11)	13.18 (0.00)	13.45 (0.15)
		4,096	13.21 (0.03)	13.18 (0.03)	13.18 (0.00)	13.12 (0.06)
		Lattice	13.18	13.18	13.18	13.18
110	40%	256	28.60 (1.27)	26.85 (1.26)	24.54 (0.02)	28.25 (0.93)
		1,024	26.78 (0.64)	25.55 (0.64)	24.47 (0.00)	26.08 (0.80)
		4,096	24.95 (0.27)	24.69 (0.24)	24.47 (0.00)	24.96 (0.31)
		Lattice	24.47	24.47	24.47	24.47

Note: Options parameter values include strike price \$100, dividend rate 10%, interest rate 5%, and time to maturity of three years. The values in parentheses are the standard error estimates based on 10 independent replications.

Table 2

Various Mesh Estimators for American Call Options with 50 Exercise Points

			Mesh Estimators			
S	σ	*b*	BG	AH	LD	MC
90	20%	256	9.96 (1.10)	7.73 (0.78)	4.54 (0.01)	8.52 (1.04)
		1,024	6.22 (0.27)	5.33 (0.29)	4.48 (0.00)	6.71 (0.41)
		4,096	5.23 (0.16)	4.94 (0.15)	4.48 (0.00)	5.66 (0.32)
		Lattice	4.47	4.47	4.47	4.47
90	40%	256	33.07 (3.75)	25.77 (2.67)	14.63 (0.02)	29.64 (4.30)
		1,024	20.75 (0.83)	17.74 (0.95)	14.43 (0.01)	22.54 (1.51)
		4,096	17.43 (0.55)	16.31 (0.56)	14.41 (0.00)	18.97 (1.17)
		Lattice	14.40	14.40	14.40	14.40
100	20%	256	16.30 (1.40)	13.13 (0.99)	8.18 (0.00)	14.12 (1.45)
		1,024	10.40 (0.32)	9.39 (0.32)	8.14 (0.00)	11.07 (0.56)
		4,096	9.09 (0.20)	8.74 (0.18)	8.14 (0.00)	9.49 (0.41)
		Lattice	8.14	8.14	8.14	8.14
100	40%	256	42.38 (4.35)	33.57 (3.05)	19.45 (0.02)	38.04 (5.15)
		1,024	26.74 (0.94)	23.55 (1.05)	19.26 (0.01)	28.92 (1.81)
		4,096	22.78 (0.66)	21.53 (0.64)	19.24 (0.00)	24.44 (1.38)
		Lattice	19.23	19.23	19.23	19.23
110	20%	256	24.14 (1.71)	20.11 (1.18)	13.44 (0.00)	21.36 (1.80)
		1,024	16.10 (0.32)	14.96 (0.26)	13.42 (0.00)	16.85 (0.67)
		4,096	14.35 (0.22)	14.00 (0.19)	13.42 (0.00)	14.83 (0.44)
		Lattice	13.42	13.42	13.42	13.42
110	40%	256	52.44 (4.93)	42.18 (3.52)	24.92 (0.01)	47.30 (5.97)
		1,024	33.32 (1.05)	29.89 (1.14)	24.76 (0.00)	35.94 (2.09)
		4,096	28.68 (0.77)	27.27 (0.71)	24.74 (0.00)	30.56 (1.57)
		Lattice	24.74	24.74	24.74	24.74

Note: Option parameters include strike price $100, dividend rate 10%, interest rate 5%, and time to maturity of three years. The values in parentheses are the standard error estimates based on 10 independent replications.

property that this method produces estimates that are high biased. The bias of the underlying method however is very high, particularly with low mesh points and/or high volatility cases. This is also supported by the larger estimates of the sample standard errors. The situation is further aggravated when we merely increase the number of exercise points from 10 to 50. For this method to be of practical use, one must incorporate extensive variance reduction techniques, as suggested by Broadie and Glasserman (1997b).

- With 10 exercise points, the average or AH mesh estimators proposed by Avramidis and Hyden (1999a) appear to be effective at reducing the bias of the BG mesh estimators. However, as we consider more frequent exercise points, this method starts to show signs of deterioration. In either case, the sample standard errors of the AH mesh estimators are comparable to that from the BG mesh estimators. This is also an indication that, although the method of Avramidis and Hyden may be effective in some cases, the estimators are subject to the same high variation as the BG mesh estimators.
- We now turn to our proposed estimators based on $\tilde{g}(t, \cdot)$ and LD methods. One immediate observation that can be drawn is that the convergence of this method is extremely high. Recall that this mesh estimator is also biased high, with the desirable property that it is asymptotically unbiased and converges to the correct value as $b \to \infty$. The results indicate that the asymptotic property can be attained even for such small mesh sizes considered in these examples. This is in contrast to the BG mesh estimators, which exhibit a much higher bias for the same number of mesh points. More intriguing, even when we increase the number of exercise points to 50, the proposed method still remains as efficient. Another important observation is that the mesh estimators based on this method are very stable, evidenced by the small sample standard errors.
- The last method, which is of expository interest only, relies on the proposed $\tilde{g}(t, \cdot)$ but with the mesh generated from the random sequences. The performance of this method is very similar to the BG mesh estimators. The poor convergence is more pronounced with a

smaller mesh size, higher volatility, and larger exercise points. Comparing this method to that based on low discrepancy sequences also leads to the conclusion that the uniformity property plays a critical role in determining the efficiency of the underlying method. When the mesh is constructed in such a way that the mesh points are simulated more uniformly from the moment-matched distribution, the efficiency of the mesh method can be greatly enhanced. This is the fundamental ingredient of the LD methods.

5.2 Multiple-Shout Options

We now consider the valuation of the embedded options in equity-linked insurance contracts. As discussed in the introduction, this type of guarantee can be formulated as a multiple shout option. The payoff function at maturity is given by

$$\max\{\max\{K, S_{\tau_1}, \ldots, S_{\tau_m}\} - S_T, 0\},$$

where K is the initial guaranteed floor level, m is the predefined number of resetting opportunities for the duration of the contract, and S_{τ_i} corresponds to the value of the fund at time τ_i and the i-th reset. Typically, the reset options embedded in the segregated funds allow two resets per year. In addition, the maturity of the contract is extended, say by another 10 years, whenever the guarantee level is reset. These additional complications are omitted in our numerical examples.

We model the shout options in the classical Black-Scholes framework. Under this assumption, Cheuk and Vorst (1997) demonstrate that these options can be priced efficiently using the trinomial lattice. For the purpose of comparison, we henceforth use the results from Cheuk and Vorst (1997) to benchmark our simulated estimates based on the various mesh methods. The parameter values for our base cases are initial fund level $S_0 = \$100$, initial guaranteed level $K = \$100$, volatility of the underlying fund $\sigma = 20\%$, annualized interest rate $r = 6\%$, and maturities of one and two years.

We consider the valuation of the m-shout options, where $m = 1, \ldots, 4$. We also assume that the shout can only occur at the end of each month, so there are 12 and 24 exercise opportunities for the contracts with one-year and two-year maturities. We consider mesh points $b \in \{256, 512, 1024, 2048, 4096\}$. For each mesh size, 10 independent meshes are generated so that the sample averages as well as their sample standard errors are calculated. Given the experience from the previous section, we only report the results for the LD and AH mesh estimators. These are presented in Figures 2 and 3. In these graphs, the estimated 90% confidence interval (from a t-distribution with 9 degrees of freedom) corresponding to each mesh size is plotted. Also shown on the graphs is the option premium (denoted by dotted horizontal line) estimated from the trinomial lattice approach.

The conclusions are very similar to those we reached in our application of the method to the valuation of American options in the previous subsection. For either contract, the LD mesh estimators reduce drastically the variability of the option estimates and exhibit a steady rate of convergence as b increases. The biases are essentially eliminated even for a relatively small mesh parameter b. This is evident from the extremely tight confidence intervals constructed. The AH mesh estimators, however, yield a much greater bias and fluctuation. The constructed confidence intervals are also much wider, compared with the LD mesh estimators.

6. CONCLUSION

Currently, the most promising procedure for the valuation of complicated American options involving several assets is the Broadie-Glasserman stochastic mesh method. In this paper, we have extended their approach to incorporate the use of low discrepancy sequences. The extended method has superior convergence properties when compared with the original Broadie-Glasserman approach. The enhanced convergence is clearly demonstrated in the valuation of the standard American options as well as in the case of multiple shout options.

It should be stressed that the underlying technique can be extended to more realistic multifactor models, and this is achieved with a very moderate increase in the computational effort. This is in contrast with other competitive techniques such as lattice or PDE approaches where the computational complexity grows exponentially with the number of state variables. When the number of state variables is greater than two or three, the

Figure 2

Mesh Estimators (and 90% Confidence Limits) for the *m*-Shout Options with $S_0 = K = \$100$, $\sigma = 20\%$, $r = 6\%$, and One Year until Expiration

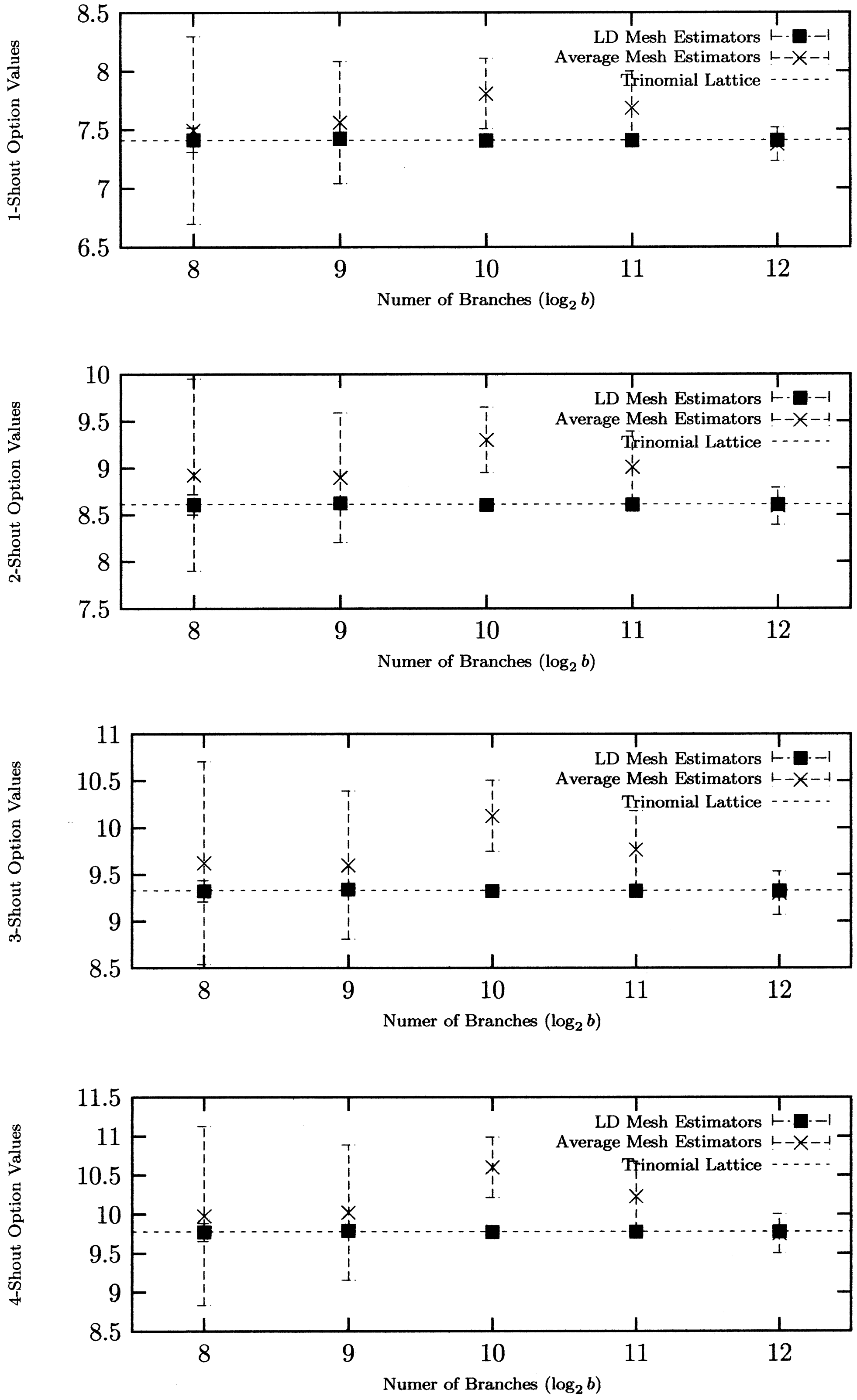
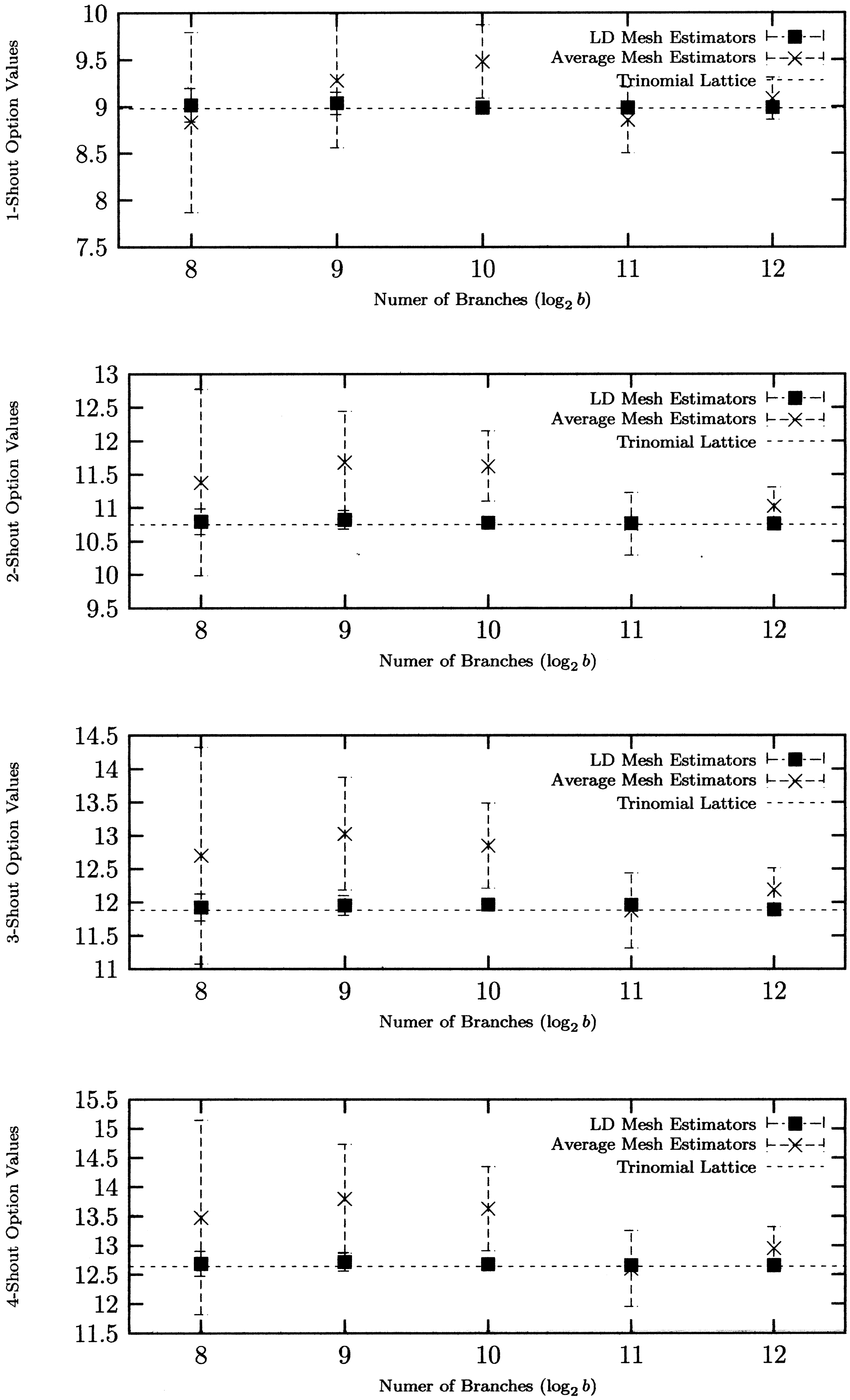


Figure 3

Mesh Estimators (and 90% Confidence Limits) for the *m*-Shout Options with $S_0 = K = \$100$, $\sigma = 20\%$, $r = 6\%$, and Two Years until Expiration



Monte Carlo methods, and in particular the proposed method, become the only viable tool. By building on some of the ideas of other researchers, we have developed a promising method for the valuation of American options. We suggest this as a particularly powerful and flexible technique not only for the valuation of reset options under segregated fund contracts, but also for the entire range of complex options that are characteristic of many of these contracts. In a related work (see Boyle, Kolkiewicz, and Tan 2000), we discuss the application of the present approach to multifactor models.

ACKNOWLEDGMENTS

Phelim P. Boyle thanks the Social Science and Humanities Research Council of Canada for research support. Adam Kolkiewicz and Ken Seng Tan acknowledge the research support from the Natural Sciences and Engineering Research Council of Canada. Phelim P. Boyle and Ken Seng Tan are also grateful to the Committee for Knowledge Extension and Research of the Society of Actuaries for financial support. They thank the participants at the *Symposium on Stochastic Modelling for Variable Annuity/Segregated Fund Investment Guarantees, Canadian Institute of Actuaries, September 13–14, 1999,* for their comments on the earlier version of this paper. The authors are grateful to two anonymous referees and the associate editor for their suggestions. This paper was previously titled "Valuation of the Reset Option in Segregated Fund Contracts using Quasi-Monte Carlo Methods."

REFERENCES

ACWORTH, P., M. BROADIE, AND P. GLASSERMAN. 1997. "A Comparison of Some Monte Carlo and Quasi-Monte Carlo Methods for Option Pricing." In *Monte Carlo and Quasi-Monte-Carlo Methods 1996,* Volume 127 of *Lecture Notes in Statistics,* ed. H. Niederreiter et al. New York: Springer-Verlag.

AVRAMIDIS, A.N., AND P. HYDEN. 1999a. "Efficiency Improvement for Pricing American Options with a Stochastic Mesh." In *Proceedings of the 1999 Winter Winter Conference,* ed. P.A. Farrington, H.B. Nembhard, D.T. Sturrock, and G.W. Evans, 344–50.

———. 1999b. "Efficiency Simulation Techniques for Pricing High-Dimensional American Options." Unpublished manuscript, Cornell University.

BARRAQUAND, J., AND D. MARTINEAU. 1995. "Numerical Valuation of High Dimensional Multivariate American Securities," *Journal of Financial and Quantitative Analysis* 30(3): 383–405.

BERMAN, L. 1997–98. "Accelerating Monte Carlo Quasi-Random Sequences and Variance Reduction," *Journal of Computational Finance* 2(2): 79–96.

BLACK, F., AND M. SCHOLES. 1973. "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy* 81(3): 637–54.

BOYLE, P.P., M. BROADIE, AND P. GLASSERMAN. 1997. "Monte Carlo Methods for Security Pricing," *Journal of Economic Dynamics and Control* 21: 1267–321.

BOYLE, P.P., A. KOLKIEWICZ, AND K.S. TAN. 2000. "Pricing American Style Options Using Low Discrepancy Mesh Methods." Technical report, IIPR 00-07, University of Waterloo.

BOYLE, P.P., AND K.S. TAN. 1997. "Quasi-Monte Carlo Methods." In *International AFIR Colloquium Proceedings, Australia,* Vol. 1, pp. 1–24.

BRATLEY, P., AND B.L. FOX. 1988. "Algorithm 659: Implementing Sobol's Quasi Random Sequence Generator," *ACM Transactions on Mathematical Software* 14: 88–100.

BRATLEY, P., B.L. FOX, AND H. NIEDERREITER. 1992. "Implementation and Tests of Low-Discrepancy Sequences," *ACM Transactions on Modeling and Computer Simulation* 2(3): 195–213.

BROADIE, M., AND P. GLASSERMAN. 1997a. "Pricing American-Style Securities Using Simulation," *Journal of Economic Dynamics and Control* 21(8–9): 1323–52.

———. 1997b. "A Stochastic Mesh Method for Pricing High-Dimensional American Options. Working paper, Columbia University.

BROADIE, M., P. GLASSERMAN, AND Z. HA. 2000. "Pricing American Options by Simulation Using a Stochastic Mesh with Optimized Weights." In *Probabilistic Constrained Optimization: Methodology and Applications,* ed. S. Uryasev, 32–50. Boston: Kluwer Academic Publishers.

CAFLISCH, R.E., W. MOROKOFF, AND A. OWEN. 1997. "Valuation of Mortgage-Backed Securities Using Brownian Bridges to Reduce Effective Dimension," *Journal of Computational Finance* 1(1): 27–46.

CARR, P., AND G. YANG. 1997. "Simulating Bermudan Interest Rate Derivatives." Working paper, Morgan Stanley/Numerix.

———. 1998. "Simulating American Bond Options in HJM Framework." Working paper, Morgan Stanley/Numerix.

CARRIERE, J.F. 1996. "Valuation of the Early-Exercise Price for Options Using Simulation and Nonparametric Regression," *Insurance: Mathematics and Economics* 19: 19–30.

CHEUK, T.H.F., AND T.C.F. VORST. 1997. "Shout Floors," *Net Exposure,* The Electronic Journal of Financial Risk, 2. Available online at http://www.netexposure.co.uk.

DEVROYE, L. 1986. *Non-Uniform Random Variate Generation.* New York: Springer-Verlag.

FAURE, H. 1982. "Discrépance de Suites Associées à un Système de Numération (en Dimension s)," *Acta Arithmetica* 41: 337–51.

FISHMAN, V., P. FITTON, AND Y. GALPERIN. 1997. "Hybrid Low-Discrepancy Sequences: Effective Path Reduction for Yield

Curve Scenario Generation," *Journal of Fixed Income* 7(1): 75.

FOX, P. 1986. "Algorithm 647: Implementation and Relative Efficiency of Quasi-Random Sequence Generators," *ACM Transactions on Mathematical Software* 12(4): 362–76.

GALANTI, S., AND A. JUNG. 1997. "Low-Discrepancy Sequences: Monte Carlo Simulation of Option Prices," *Journal of Derivatives* 5(1): 63–83.

GRANT, D., G. VORA, AND D. WEEKS. 1997. "Path-Dependent Options: Extending the Monte Carlo Simulation Approach," *Management Science* 43(11): 1589–1602.

HALTON, J.H. 1960. "On the Efficiency of Certain Quasi-Random Sequences of Points in Evaluating Multi-Dimensional Integrals," *Numerische Mathematik* 2: 84–90.

IBANEZ, A., AND F. ZAPATERO. 1999. "Monte Carlo Valuation of American Options Through Computation of the Optimal Exercise Frontier." Working Paper, Instituto Tecnologico Autonomo de Mexico.

JOY, C., P.P. BOYLE, AND K.S. TAN. 1996. "Quasi-Monte Carlo Methods in Numerical Finance," *Management Science* 42(6): 926–38.

MITCHELL, G.T. 1998. "Improving Monte Carlo Modeling with Low Discrepancy Sequences—a Test," *The Financial Reporter (Society of Actuaries)* 36: 19–22.

NIEDERREITER, H. 1987. "Point Sets and Sequences with Small Discrepancy," *Monatshefte fur Mathematik* 104: 273–337.

———. 1992. *Random Number Generation and Quasi-Monte Carlo Methods.* Philadelphia: S.I.A.M.

NINOMIYA, S., AND S. TEZUKA. 1996. "Toward Real-Time Pricing of Complex Financial Derivatives," *Applied Mathematical Finance* 3: 1–20.

OWEN, A.B. 1996. "Randomly Permuted (t, m, s)-Nets and (t, s)-Sequences." In *Monte Carlo and Quasi-Monte-Carlo Methods in Scientific Computing: Proceedings of a Conference at the University of Nevada, Las Vegas, June 23–25,* ed. H. Niederreiter and P.J-S. Shiue, pp. 299–317.

OWEN, A.B., AND D. TAVELLA. 1997. "Scrambled Nets for Value at Risk Calculations." In *VaR: Understanding and Applying Value-at-Risk,* ed. S. Grayling. London: RISK.

PASKOV, S.H. 1996. "New Methodologies for Valuing Derivatives." In *Mathematics of Derivative Securities,* ed. S. Pliska and M. Dempster. Cambridge, England: Cambridge University Press, Isaac Newton Institute.

PASKOV, S.H., AND J.F. TRAUB. 1995. "Faster Valuation of Financial Derivatives," *Journal of Portfolio Management* 22(1): 113–20.

PRESS, W., S. TEUKOLSKY, T.V. WILLIAM, AND P.F. BRIAN. 1992. *Numerical Recipes in C.* 2nd ed. New York: Cambridge University Press.

RAYMAR, S., AND M. ZWECHER. 1997. "A Monte Carlo Valuation of American Options on the Maximum of Several Stocks," *Journal of Derivatives* 5(1): 7–23.

SHUKHMAN, B.V. 1994. "Generation of Quasi-Random (lp_τ) Vectors for Parallel Computation," *Computer Physics Communications* 78: 279–86.

SOBOL', I.M. 1967. "The Distribution of Points in a Cube and the Approximate Evaluation of Integrals," *U.S.S.R Computational Mathematics and Mathematical Physics* 7(4): 86–112.

TAN, K.S. 1998. "Quasi-Monte Carlo Methods: Applications in Finance and Actuarial Science." Ph.D. thesis, University of Waterloo, Waterloo, Canada.

TAN, K.S., AND P.P. BOYLE. 1997. "Applications of Scrambled Low Discrepancy Sequences to Exotic Options," *International AFIR Colloquium Proceedings, Australia,* Vol. 2, pp. 885–918.

———. 2000. "Applications of Randomized Low Discrepancy Sequences to the Valuation of Complex Securities," *Journal of Economic Dynamics and Control* 24: 1747–82.

TEZUKA, S. 1993. "Polynomial Arithmetic Analogue of Halton Sequences," *ACM Transactions on Modeling and Computer Simulation* 3: 99–107.

THOMAS, B. 1993. "Something to Shout About," *RISK* 6: 56–58.

TILLEY, J.A. 1993. "Valuing American Options in a Path Simulation Model," *Transactions of the Society of Actuaries* XLV: 499–520.

WINDCLIFF, H., P.A. FORSYTH, AND K.R. VETZAL. 1999. "Shout Options." Working Paper, University of Waterloo.

Discussions on this paper can be submitted until Jan. 1, 2002. The authors reserve the right to reply to any discussion. Please see the Submission Guidelines for Authors on the inside back cover for instructions on the submission of discussions.