A Simple Approach to the Pricing of Bermudan Swaptions in the Multi-Factor Libor Market Model

Leif Andersen[1]

General Re Financial Products
630 Fifth Avenue, Suite 450
New York, NY 10111

First Version: November 23, 1998
This Version: March 5, 1999

Abstract

This paper considers the pricing of Bermuda-style swaptions in the *Libor market model* (Brace *et al* (1997), Jamshidian (1997), Miltersen *et al* (1997)) and its extensions (Andersen and Andreasen (1998)). Due to its large number of state variables, application of lattice methods to this model class is generally not feasible, and we instead focus on a simple technique to incorporate early exercise features into the Monte Carlo method. Our approach involves a direct search for an early exercise boundary parametrized in intrinsic value and the values of still-alive swaptions. We compare results of the proposed algorithm against prices obtained from Markov Chain approximations and finite difference methods. The proposed algorithm is fast and robust, and produces a lower bound on Bermudan swaption prices that appears to be very tight for many realistic structures. The paper contains several numerical results against which other methods can be tested.

1. Introduction.

A *Bermudan swaption* is an option which at each date in a schedule of exercise dates gives the holder the right to enter into an interest rate swap, provided this right has not been exercised at any previous time in the schedule. Due to their usefulness as hedges for callable bonds, Bermudan swaptions are actively traded and probably the most liquid fixed income instrument with a built-in early exercise feature. To compute prices of these instruments, most banks choose to use simple models that involve only one or two state variables. Commonly applied models include the one-factor short-rate models of Black *et al* (BDT) (1990), Black and Karasinski (1991), Hull and White (1990); and Ritchken and Sankarasubramanian (1995), just to name a few. A common trait of all these models is the fact that they can be implemented numerically in low-dimensional *lattices* (such as finite differences or binomial trees) which are well-suited for dealing with the free boundary condition that arises for options with early exercise features. The ease with which the models above can handle American- and Bermudan-style instruments, however, comes at the expense of realism. For instance, all models above use only one driving Brownian motion and as such imply perfect correlation of all forward rates. Also, by working solely with the short

---

[1] I am grateful to Mark Broadie, Guang Yang, and Morten B. Pedersen for helpful conversations. I also wish to thank my colleagues at GRFP, in particular Dirk Bangert whose experiments with equity options inspired the approach taken in this paper, and Jesper Andreasen who produced the BDT results reported in Section 4.

rate as the model primitive, the degrees of freedom available in the model are too low to allow for a precise fit to prices of quoted instruments (caps and swaptions). Moreover, if such a fit is attempted by making, say, the short-rate volatility time-dependent, commonly the evolution of the term structure of volatilities becomes non-stationary and largely unpredictable (see e.g. Carverhill (1995) and Hull and White (1995) for a discussion of this issue).

To introduce a more realistic model framework, one can, for instance, turn to the so-called *Libor market (LM) model* (Brace *et al* (1997), Jamshidian (1997), Miltersen *et al* (1997)) and its extensions (Andersen and Andreasen (1998)). This model class readily allows for multiple stochastic factors, can incorporate volatility skews, prices liquid market instruments in closed form, and has enough degrees of freedom to allow for a good fit to market prices of caps and swaptions while still maintaining a largely stationary volatility term structure. Not surprisingly, all these desirable features come with a price tag: due to very high number of state variables in the LM model (often more than 50), recombining lattices are not computationally feasible and pricing of contingent claims must virtually always be done by *Monte Carlo simulation*. Although flexible and easy to implement, the Monte Carlo method has several drawbacks, including slow convergence and difficulties in dealing with derivatives that contain early exercise features (e.g. Bermuda-style swaptions). The first of these two problems can normally be handled by application of one or more so-called variance reduction techniques (many of which are surveyed in the excellent review article by Boyle *et al* (1997)); indeed, it is fair to say that most option computations in the LM model can be set up in such a way that computation times are at least acceptable. Coping with early exercise features is thornier, and, in fact, until recently was widely considered beyond the limitations of the Monte Carlo method. Recent work by Tilley (1993), Broadie and Glasserman (1997a), and others, however, have proven this belief to be incorrect, although the practical obstacles still remain rather formidable.

Broadly speaking, three different approaches to the pricing of American-style derivatives in the LM model (or the closely related HJM model by Heath *et al* (1992)) have been proposed in the literature. In the first approach, a non-recombining tree (sometimes known as a "bushy" tree) is set up to approximate the continuous-time dynamics of interest rates (see e.g. Heath *et al* (1990) or Gatarek (1996) for details). Backwards induction algorithms can then be applied and the early exercise feature easily incorporated. Unfortunately, the number of nodes in the bushy tree grows exponentially in the number of its time-steps: for m stochastic factors and n time-steps, the total number of nodes equals $m^{-1}[(m+1)^{n+1}-1]$. So, for just 15 time-steps in a 3-factor model, the total number of tree nodes would equal around 1.4 billion (!). As many long-dated derivatives require significantly more than 15 time-steps to achieve convergence, bushy trees are far too slow for general pricing, although sometimes useful for short-dated instruments. As an aside, we point out that Broadie and Glasserman (1997a) devise a Monte Carlo method for Bermuda-style options around the concept of simulated non-recombining trees. Again, severe restrictions on the feasible number of time-steps apply[2].

[2]In a more recent paper, Broadie and Glasserman (1997b) also devise a Monte Carlo technique around a stochastic recombining lattice, a so-called *stochastic mesh*. This method poses its own challenges and, as evident from the recent work by Pedersen (1999), is not easily adapted to the LM model framework.

A second method, exemplified by Carr and Yang (1997), is based on the stratification technique of Barraquand and Martineau (1995). Briefly, Carr and Yang set up buckets for the money market account in the LM model and enforce Markov chain dynamics on the transition between buckets at different time-steps. The transition probabilities of the Markov chain are constructed empirically through Monte Carlo simulation. During this simulation, each particular bucket of the money market numeraire is associated with a state of the yield curve; this curve is found by averaging all simulated yield curves that passed through the bucket. Once the Markov chain of numeraires and yield curves has been constructed, pricing of American and Bermudan options can be found by a backwards induction algorithm similar to the one applied in lattices. While the approach of Carr and Yang seems to give good results in many cases (as we shall see later), the chosen stratification variable (the money market account) is only a weak indicator of the state of the yield curve[3]. Moreover, it appears very difficult to analyze the errors and biases associated with the Markov chain approach. The algorithm is thus affected by various sources of potential biases, including a) a bias from forcing Markovian dynamics on a non-Markov variable; b) a bias from averaging yield curves at each bucket; c) a bias from basing the exercise decision solely on the state of the numeraire; and d) a bias from, in effect, using the same random paths to determine both the exercise strategy and the option price. The bias in c) is negative (suboptimal exercise strategy), whereas the bias in d) is positive; the effects of a) and b) are more difficult to predict and may depend on the specific option payout. For instance, for an option on the spread between a long and a short rate, averaging yield curves would normally tend to dampen curve steepenings; consequently, for this particular option, the bias in b) would likely be negative[4].

A third method for handling Bermudan swaptions has been suggested in Clewlow and Strickland (1998). As in Carr and Yang (1997), their method is based on reducing the exercise decision to the state of a single variable, in their case the value of the fixed side of the underlying swap. Working in a two-factor Gaussian HJM model, Clewlow and Strickland determine the early exercise boundary by extracting information from a best-fit one-factor Gaussian model implemented in a lattice. This boundary is then used in a Monte Carlo simulation of the full two-factor model. Clewlow and Strickland's method is simple and robust, but will only return a lower bound on the price (due to the suboptimal exercise policy). As the information generated from a one-factor model might be of limited value in a multi-factor setting, it is not inconceivable that this bias could be large in some circumstances. Nevertheless, the fact that the bias of this method has a predictable sign allows for better control of model risk than the method of Carr and Yang.

The method proposed in this paper is similar in spirit to that of Clewlow and Strickland (1998) in that we attempt to parametrize the early exercise boundary in the state of a very few variables, primarily the intrinsic value of the underlying swap. As such, our method will also

[3] The money market account is, essentially, an accumulator of the state of interest rates through time. As such, the information contained in the money market account about current interest rates is very limited. For instance, a path of high rates followed by low rates may yield the same value of the money market account as a path of low rates followed by high rates.

[4] Carr and Yang (1997) realize the potential problem of averaging yield curves for options that depend heavily on curve steepenings. In this situation, they advocate keeping track of the average option payoff rather than the average yield curve.

generate a lower bound on the true option price. Unlike Clewlow and Strickland, however, the early exercise boundary is not determined in a one-factor lattice, but found by optimization on the results of a separate simulation of the full multi-factor model. This technique is somewhat similar to the equity option algorithms discussed in Grant *et al* (1997) and Li (1996), and can conveniently be decomposed into a recursive series of simple one-dimensional optimization problems. As the dependence of most Bermudan swaption prices on the exact location of the early exercise boundary turns out to be quite weak, the process of locating the exercise barrier can normally be set up to be very fast, typically much faster than the subsequent pricing of the Bermudan swaption through Monte Carlo simulation. In general, the proposed algorithm is only slightly slower than pricing a regular European swaption maturing at the last exercise date of the Bermudan swaption.

The rest of this paper is organized as follows: in Section 2 we introduce the LM model in the "extended" form of Andersen and Andreasen (1998) and define the payout function for a Bermudan swaption. In section 3, we discuss our approach to pricing Bermudan swaption and discuss various approaches to parametrization of the early exercise boundary. Section 4 contains numerical results for the one-factor case and compares the results of our method with those of a best-fit BDT short-rate model. The section examines the robustness of the proposed method, and also briefly looks at the effects of changing volatility skews. Section 5 considers the multi-factor case and contains comparisons with published results based on bushy trees and Markov chains. Finally, section 6 contains the conclusions of the paper.

2. Notation and General Framework.

The model used in this paper is the "extended" LM model of Andersen and Andreasen (1998), building upon the log-normal model of Brace *et al* (1997), Jamshidian (1997), and Miltersen *et al* (1997). This section will contain a brief summary of this model; for a detailed discussion see Andersen and Andreasen (1998).

Consider an increasing maturity structure $0 = T_0 < T_1 < ... < T_{K+1}$ and define a right-continuous mapping function $n(t)$ by

$$T_{n(t)-1} \le t < T_{n(t)} .$$

To match conventions in swap and cap markets, we would often use a nearly equidistant spacing between points (say 3 or 6 calendar months) in the maturity structure. With $P(t,T)$ denoting the time *t* price of a zero-coupon bond maturing at time *T*, we define discrete forward rates (*Libor forward rates*) on the maturity structure as follows:

$$F_k(t) \equiv \frac{1}{d_k}\left(\frac{P(t,T_k)}{P(t,T_{k+1})} - 1\right), \quad d_k = T_{k+1} - T_k ,$$

or

$$P(t,T_k) = P(t,T_{n(t)}) \prod_{j=n(t)}^{k-1} (1+ d_j F_j(t))^{-1} .$$

For this definition to be meaningful, we must require that $t \le T_k$ and $k \le K$. For brevity, we will omit such obvious restrictions on time and indices in most of the equations that follow.

We now introduce a discrete-time Libor money market account B as the current value of the strategy of "rolling over" an initial investment of \$1 at each time in the maturity structure. Specifically,

$$B(t) = P(t,T_{n(t)}) \prod_{j=0}^{n(t)-1} P(T_j,T_{j+1})^{-1} = P(t,T_{n(t)}) \prod_{j=0}^{n(t)-1} (1+ d_j F_j(T_j)) \qquad . \tag{1}$$

The process $B(t)$ in (1) is positive and can be used as a pricing numeraire. The probability measure Q induced by this choice of numeraire is normally called the *spot measure* and is closely related to the usual risk-neutral measure induced by a continuously rolled money market account. Under the extended LM model, the no-arbitrage dynamics of forward rates are governed by the following set of stochastic differential equations, $k = n(t),\ldots,K$:

$$dF_k(t) = j(F_k(t)) l_k(t)^{\mathrm{T}} \left[m_k(t)dt + dW(t) \right], \quad m_k(t) = \sum_{j=n(t)}^{k} \frac{d_j j(F_j(t)) l_j(t)}{1+ d_j F_j(t)} . \tag{2}$$

In (2), $j : \Re^+ \to \Re^+$ is a one-dimensional function satisfying certain regularity conditions (see Andersen and Andreasen (1998)), $l_k(t)$ is a bounded m-dimensional deterministic function, and $W(t)$ is a m-dimensional Brownian motion under Q. (2) defines a system of up to K Markov state variables.

For any sufficiently regular choice of j, Andersen and Andreasen (1998) demonstrate how caps and European swaptions can be priced efficiently in a small set of finite difference grids, enabling fast calibration of the $l_k(t)$ functions to market data. The so-called *CEV model* sets j to a power function, $j(x) = x^a$, $a > 0$, and is analytically tractable. In particular, caps can be priced in closed-form, and excellent closed-form approximations exist for European swaptions (see Andersen and Andreasen (1998) for details). Setting $a < 1$ will generate a downward sloping volatility skew[5]; $a > 1$ an upward-sloping one.

To illustrate how the LM model is used for derivatives prices, consider the pricing of a European claim V with the terminal payout $V(T) = g(T)$ at some future maturity $0 \le T \le T_{K+1}$. The payout function $g(T)$ is allowed to depend on the entire path[6] of the forward curve from 0 to T. By standard arguments, the arbitrage-free price at time 0 of this claim can be written

---

[5] That is, caplet volatilities will be a downward sloping function of strike.

[6] Technically, g is thus simply required to be a stochastic process adapted to the information (filtration) generated by the vector-valued Brownian motion W up to time T

$$V(0) = E^{Q}\left[g(T)/B(T)\right], \qquad (3)$$

where $E^{Q}[\cdot]$ denotes expectation under Q.

Due to the high number of state-variables necessary to describe the path of the yield curve, (3) normally must be evaluated through Monte Carlo simulation. That is, we generate a large number of random paths of the forward curve and evaluate the expectation in (3) as a simple average over the random paths. Schemes to discretize (2) into a form suitable for Monte Carlo simulation can be found in Andersen and Andreasen (1998); the basics of Monte Carlo methods are discussed in Boyle *et al* (1997). All Monte Carlo simulations in this paper were based on a first-order log-Euler discretization with a time-step equal to the frequency of the forward rates[7].

As an example of a European interest rate option consider a *European swaption* $S_{s,e}$ maturing at some date T_s, $s \in \{1,2,\ldots,K\}$. The swaption gives the holder the right to enter into a fixed-floating interest rate swap where fixed cashflows $q\, d_{k-1} > 0$ paid at T_k, $k = s+1, s+2, \ldots, e$ are swapped against floating Libor (paid in arrears) on a \$1 notional. T_s and T_e are thus the start- and end-dates of the underlying swap, respectively, and clearly we require $T_{K+1} \geq T_e > T_s$. Notice that we only consider swaps with cash-flow dates that coincide with the maturity structure. At maturity T_s the value of $S_{s,e}$ is, by definition,

$$S_{s,e}(T_s) = \left(f \sum_{k=s}^{e-1} P(T_s, T_{k+1})\, d_k [F_k(T_s) - q]\right)^{+} = \left(f[1 - P(T_s, T_e)] - fq \sum_{k=s}^{e-1} d_k P(T_s, T_{k+1})\right)^{+}. \qquad (4)$$

In (4), the flag f is +1 if the option holder has the right to pay fixed and receive floating (payer swaption), and -1 if the option holder receives fixed and pays floating (receiver swaption). The time 0 value of $S_{s,e}$ can, as mentioned earlier, be approximated either in closed form or by the solution to a one-factor finite difference grid (see Andersen and Andreasen (1998) for details).

In the American or Bermudan version of the general European claim in (3), the option holder is further granted the right to exercise the option early. Let us denote the (random) time of early exercise t, $0 \leq t \leq T$, that is, the option holder receives receive $g(t)$ at time t. Since the option holder rationally should choose t to optimize the value of the option, we can write for the American/Bermudan option value $\overline{V}$:

$$\overline{V}(0) = \sup_{t \in \aleph} E^{Q}[g(t)/B(t)]. \qquad (5)$$

where $\aleph$ denotes the set of all allowed exercise strategies[8]. For an American option, $\aleph$ would be valued in $[0,T]$; for a Bermudan option, $\aleph$ would be valued in some discrete set of exercise

[7] As an exception, the numbers in Tables 6a-b were generated using a time-step equal to one-quarter of the forward rate frequency. This was done mainly for consistency with numbers generated by Carr and Yang (1997) and typically affected prices by less than 1 or 2 basis points.

[8] Technically, these exercise times are stopping times adapted to the filtration generated by the path of *W*. Similarly, the function *g* is also assumed an adapted process on the filtration generated by *W*.

times, for instance $\{0, T_1, T_2, \ldots, T_{n(T)-1}, T\}$. Let us use t^* to denote the (optimal) early exercise strategy that solves (5). Assuming that t^* can be computed one way or the other, (5) can be evaluated by straightforward Monte Carlo simulation. As discussed in the previous section, however, determination of the optimal exercise strategy in the LM model is highly non-trivial.

In this paper, we will focus on a particular type of Bermudan option, namely the *Bermudan swaption*. For our purposes, a Bermudan swaption $\bar{S}_{s,x,e}$ is characterized by three dates: the lock-out date (T_s), the last exercise date (T_x), and the final swap maturity (T_e). We assume that $T_s < T_x < T_e$, and that all three dates coincide with dates in the maturity structure; that is, *s*, *x*, and *e* are all integers in {0,…,*K*+1}. Early exercise of the Bermudan swaption is restricted to dates in the discrete set $\{T_s, T_{s+1,} \ldots, T_x\}$. Assuming that exercise takes place at, say, $t = T_i$, the option holder receives, at time T_i,

$$\bar{S}_{s,x,e}(T_i) = S_{i,e}(T_i) \tag{6}$$

where $S_{i,e}$ is the European swaption defined in (4). We point out that virtually all swaptions traded in the market have $T_s > 0$ and $T_x = T_{e-1}$. We have allowed for a more flexible definition of Bermudan swaptions mainly to be able to compare against certain results in the literature[9].

3. Monte Carlo simulation of Bermudan Swaptions.

In this section, we will discuss an approximate approach to determining the early exercise strategy for a Bermudan swaption. To this end, let us first introduce an early exercise indicator function $I(t)$ that equals 1 if early exercise is optimal at time *t*, and 0 otherwise. That is, with the notation introduced earlier,

$$t^* = \inf\{t \in \{T_s, T_{s+1}, \ldots, T_x\}: I(t) = 1\}.$$

For the Markov system (2), the decision of whether or not to exercise a Bermudan swaption on some time $T_i \in \{T_s, T_{s+1,} \ldots, T_x\}$ will generally depend in a complicated way on the state of all forward rates $F_k(T_i)$, $k = i, i+1, \ldots, e-1$. As this "correct" exercise strategy depends on too many state variables to be feasible in a Monte Carlo setting, we first attempt to reduce the dimensionality of the exercise decision by postulating the following form of $I(t)$:

$$I(T_i) \approx f\left(S_{i,e}(T_i),\ S_{i+1,e}(T_i),\ S_{i+2,e}(T_i), \ldots, S_{x,e}(T_i);\ H(T_i)\right), \tag{7}$$

where f is some specified Boolean function with a single, possibly time-dependent, parameter $H(\cdot)$. That is, we assume that the exercise decision depends solely on the European values of all

---

[9] Some authors consider another variety of Bermudan swaptions where the number of cash-flows in the underlying swap is independent of the time of exercise (that is, $T_e = t + c$ for some constant *c*). These so-called *constant maturity* Bermudan swaptions are rarely traded in practice and will not be considered in this paper. We do point out, however, that these structures are easily priced in the framework developed in this paper.

still-alive "component" swaptions of the Bermudan swaption. The form (7) is convenient in the LM model which, as mentioned earlier, allows for efficient pricing of European swaptions. Moreover, it appears reasonable to assume that the decision to exercise a Bermudan swaption depends strongly on the absolute and/or relative values of its underlying European options.

Common for all specifications of f is the boundary condition

$$I(T_x) = \begin{cases} 1, & if\ S_{x,e}(T_x) > 0 \\ 0, & otherwise \end{cases}, \tag{8}$$

which just states that the Bermudan swaption will be exercised at T_x (the last possible date) if and only if the underlying swap is in-the-money at that date. One can imagine many different reasonable specifications of the function *f* in (7) that satisfy (8). Below, we list two simple suggestions which that we will examine closer in the following. As the algorithms we present later are independent of the exact form of *f*, the reader should feel free to come up with other, possibly better, specifications[10] of *f*.

<u>Approximate exercise strategy I:</u>

$$I(T_i) = \begin{cases} 1, & if\ \ S_{i,e}(T_i) > H(T_i) \\ 0, & otherwise \end{cases} \tag{9}$$

<u>Approximate exercise strategy II:</u>

$$I(T_i) = \begin{cases} 1, & if\ \ S_{i,e}(T_i) > H(T_i)\ and\ \ \underset{j=i+1,\ldots,x}{MAX}\left(S_{j,e}(T_i)\right) \le S_{i,e}(T_i) \\ 0, & otherwise \end{cases} \tag{10}$$

In strategy I, exercise takes place when, in effect, the intrinsic value of the underlying swap exceeds some time-varying barrier H. The second strategy is a refinement that also checks whether one or more of the remaining European swaptions has a value that exceeds the intrinsic swap value. If this is the case, strategy II decides that exercise cannot be optimal -- a reflection of the fact that a Bermudan swaption can always be sold at the value of its most expensive European component swaption. In general, strategy II can be expected to be most useful in a multi-factor model where the correlation of the European swaptions underlying the Bermudan structure is lower than 1. By (8), in both strategies I and II we have $H(T_x) = 0$.

With (7) and some specified form of the function *f*, determination of the early exercise strategy is solely a matter of determining the deterministic function $H(t)$ for all

---

[10] Inspired by Carr and Yang (1998), we also experimented with exercise strategies based on the level of the short rate. Such strategies, however, generally appear to give rise to lower Bermudan price estimates than strategies based on intrinsic value, particularly (and not surprisingly) in a multi-factor setting.

$t \in \{T_s, T_{s+1}, \ldots, T_x\}$. $H(t)$ is characterized by being the function that maximizes the value of the Bermudan swaption price given the exercise criterion (7). One naïve way of locating $H(\cdot)$ might thus be to execute a brute-force search for the $x-s$ levels $H(T_s), H(T_{s+1}), \ldots, H(T_{x-1})$ that maximize the price of the Bermudan swaption subject to the chosen form of the exercise strategy (7). In each iteration of the resulting high-dimensional optimization problem, the exercise strategy would be known explicitly and we could use Monte Carlo simulation to determine the price of the option. Such an approach, however, would be hopelessly slow.

To find an alternative to brute-force search for the function $H(\cdot)$, consider the Bermudan swaption $S_{x-1,x,e}$ that starts at the second-to-last exercise date T_{x-1}. At T_{x-1}, the optimal early exercise strategy for $\bar{S}_{x-1,x,e}$ must clearly be the same as for $\bar{S}_{s,x,e}$ where $s < x-1$ (since the two Bermudan swaptions are identical at time T_{x-1}). Given (8), the value of $H(T_x)$ is known and finding the value of $H(T_{x-1})$ reduces to a *one*-dimensional optimization procedure to maximize the value of $\bar{S}_{x-1,x,e}(T_{x-1})$. For each guess for $H(T_{x-1})$ in this procedure, the price of $\bar{S}_{x-1,x,e}(T_{x-1})$ can be determined by Monte Carlo simulation. Once $H(T_{x-1})$ is found this way, we can repeat the one-dimensional optimization procedure for $\bar{S}_{x-2,x,e}(T_{x-2})$ to determine $H(T_{x-2})$, and so forth all the way back to $H(T_s)$.

In the backward-starting algorithm above, notice that it is not necessary to repeat the Monte Carlo simulation for each iteration in the $x-s$ sequential one-dimensional optimization procedures. Instead, the numbers from a single Monte Carlo simulation session can be stored in memory and used over and over. The memory requirements depend on the specific form of f but are generally relatively modest. For instance, for the strategy I in (9), each simulated path requires 2 numbers -- intrinsic swap value and the spot numeraire B -- to be stored for each date $T_s, T_{s+1}, \ldots, T_x$. With this strategy, n Monte Carlo paths requires storage of a total of $2(x-s+1)n$ floating point numbers[11]. Even for a long-dated Bermudan swaption with, say, 20 exercise dates, 16MB of memory would be sufficient to store around 50,000 Monte Carlo paths. As we shall see later, the number of simulations necessary to determine $H(\cdot)$ to adequate precision is normally much smaller than this.

Once the function values $H(T_s), H(T_{s+1}), \ldots, H(T_{x-1})$ have been found, we discard all stored Monte Carlo paths and run a *separate* Monte Carlo simulation to determine the price of the Bermudan swaption. That is, we simulate the price of the Bermudan swaption using the approximate specification (7) where the necessary values of $H(\cdot)$ are now all known. To keep the statistical error on the price estimate low, this second simulation session would normally use many more paths than the session used to determine the exercise boundary. To avoid any perfect foresight bias (a bias that is introduced by using the same random numbers to determine the exercise strategy and the price), notice that the random numbers used in the two simulations must be independent. This way, we are ensured that the only bias that affects the computed price is the

[11] Strategy II in (10) requires storage of around $3(x-s+1)N$ floating point numbers since we also need to store the largest European swaption values.

one originating from the sub-optimal choice of exercise boundary. In other words, we are guaranteed that our method generates a true lower bound[12].

To summarize, here is the proposed algorithm to price a Bermudan swaption $\bar{S}_{l,x,e}(0)$:

Step 1: Decide on a functional form f for the exercise strategy in (7). f should depend on the values of European swaptions and one time-dependent function $H(t)$.

Step 2: Run an n-path Monte Carlo simulation where for each path and each time $T_s, T_{s+1}, \ldots, T_x$ the following is stored in memory: i) intrinsic value; ii) the numeraire B; iii) other data necessary to compute f. (For strategy II in (10), say, iii) would be the maximum value of the remaining European swaptions, see footnote 11).

Step 3: Using (5), (7), and the numbers stored in i) and ii) in Step 2, compute the values $H(T_s),\ H(T_{s+1}), \ldots, H(T_{x-1})$ such that the value of the Bermudan swaption is maximized. This optimization problem can be done in backwards fashion starting with $H(T_{x-1})$ and the boundary condition (8). In total, $x-s$ simple one-variable optimization[13] problems need to be solved to determine the exercise strategy. (For improvements to this step, see the discussion below).

Step 4: Using the exercise strategy found in Step 3, price the option by an independent N-path (N>>n) Monte Carlo simulation of (5).

As mentioned earlier, the dependence of a Bermudan swaption of the exact location of the exercise barrier is often quite weak. Frequently, one can take advantage of this by guessing on a form of $H(\cdot)$ that involves only a few parameters. For instance, for long-dated instruments it is often very useful to assume that $H(\cdot)$ can be approximated by a piecewise linear function with fewer kink-points q than the number of exercise dates $(x-s+1)$ in the swaption. Since values of $H(\cdot)$ between kink-points can be computed by linear interpolation, the number of one-dimensional optimization problems to be solved in Step 3 above will be reduced to the chosen number of kink-points $q < (x-s+1)$. In general, the fewer parameters that have to be estimated in Step 3, the less Monte Carlo simulations (n) are necessary to get a smooth, noise free estimation of the exercise boundary.

As a final comment, we point out that the algorithm above can be extended to functions f that depend on more than one free parameter. This obviously complicates the optimization procedure somewhat, although it will likely remain manageable for a few (say 2 or 3) free parameters. Also, while the form (7) is specific to Bermudan swaptions, we can generalize to

[12] The idea of avoiding perfect foresight biases by separating the simulations that determine exercise rule and option price was originally suggested by Mark Broadie; see e.g. Raymar and Zwecher (1997), p. 21, footnote 9.

[13] Examples of well-known one-dimensional optimization algorithms include Golden Section Search and Brent's method, both described in detail in Press *et al* (1992)

other securities and to functions f that depend on the yield curve and/or its path in almost arbitrarily complicated manners (as long as the number of free parameters is manageable).

4. Numerical results for the one-factor model.

In this and the following section we will compare the results of our algorithm with those produced by other techniques. In doing so, we generally keep the parametrization of the LM model very simple (flat Libor forward curve, parametric volatility term structure, etc.). This serves the purpose of making our results easily reproducible and also simplifies the comparisons with models that incorporate term structures of forwards and volatilities through different mechanisms than the LM model[14]. We have, of course, tested our algorithms on models calibrated to real-life data. The results of these tests were generally consistent with the results reported here.

Some of the results reported in this section are used in a recent paper by Pedersen (1999) in a comparison of the approach in this paper with the technique of Broadie and Glasserman (1997b) and a regression-based method introduced by Longstaff and Schwarz (1998).

4.1. Comparison with the BDT one-factor short-rate model.

As a first test-case, we consider a flat semi-annual (i.e. $d_k = 0.5$ for all k) forward curve at 6%, the dynamics of which are driven by a one-factor log-normal LM model (i.e. m = 1, $j(x) = x$) with constant volatility. We are interested in pricing standard at-the-money (ATM) Bermudan swaptions for which the fixed coupon equals $q = 6\%$ and where early exercise can take place at all coupon dates following the initial lock-out period (i.e. $T_x = T_{e-1}$). To test our results, we compare with the prices obtained in a BDT short-rate model with a constant short-rate volatility s. The appropriate value of s is determined by fitting the BDT model to the appropriate European swaption prices $S_{s,e}(0), S_{s+1,e}(0), \ldots, S_{x,e}(0)$ generated by the LM model.

In Table 1 below, we list various European and Bermudan swaption prices produced by a 200 x 200 BDT Crank-Nicholson finite difference lattice as well as by Monte Carlo simulation of the corresponding LM model. The volatilities in the table were picked to be roughly consistent with current USD market conditions. For the Bermudan options, determination of the early exercise boundary in the LM model was done using $n = 10{,}000$ Monte Carlo simulations on Strategy I in (9); the subsequent pricing of the Bermudan swaption was done using $N = 50{,}000$ simulations with antithetic sampling. Numbers in parenthesis denote sample standard deviations.

In Table 1, we first notice that the European swaption prices of the BDT and LM models are very close to each other, suggesting that the dynamics of the two models are similar. For the short- and medium-dated Bermudan swaptions, the prices generated in the LM model by our Monte Carlo algorithm are essentially indistinguishable from those of generated by the BDT lattice. For the 20-year swaption with 10-year lockout, however, it appears that the results of the LM model are biased low, around 3 basis points (or less than 1% of total value), relative to the BDT model. It is, of course, difficult to break down this slight bias into a term stemming from

[14] As we discussed earlier, a short-rate model, for instance, fitted to a non-flat volatility term structure will normally imply a different evolution of spot interest rate volatility than will the LM model.

differences in model dynamics[15] and a term stemming from the inherent suboptimality of the exercise boundary used in the LM model. In general, the results of Table 1 are encouraging.

Swaption Prices (Basis Points) in BDT and 1-Factor LM Models (Strategy I)
$d_k = 0.5$; $F_k(0) = 6\%$; $q = 6\%$; $j(x) = x$; $T_x = T_{e-1}$; $N = 50{,}000$; $n = 10{,}000$

T_s	T_e	(B)ermudan (E)uropean	f	LM l_k (const.)	BDT Best Fit s	LM Price (SD)	BDT Price
1	4	E	+1	20%	20.12%	121.9 (0.5)	122.1
2	4	E	+1	20%	20.12%	111.2 (0.5)	111.1
3	4	E	+1	20%	20.12%	66.0 (0.3)	65.9
1	4	B	+1	20%	20.12%	157.7 (0.5)	158.0
1	4	B	-1	20%	20.12%	156.6 (0.3)	156.8
2	5	E	+1	20%	20.12%	162.0 (0.7)	162.2
3	5	E	+1	20%	20.12%	128.2 (0.6)	128.2
4	5	E	+1	20%	20.12%	71.7 (0.3)	71.7
2	5	B	+1	20%	20.12%	187.9 (0.6)	187.8
2	5	B	-1	20%	20.12%	186.6 (0.4)	186.8
5	10	E	+1	15%	15.22%	252.3 (1.0)	252.4
6	10	E	+1	15%	15.22%	214.6 (0.8)	213.8
7	10	E	+1	15%	15.22%	168.6 (0.7)	168.0
8	10	E	+1	15%	15.22%	116.5 (0.5)	116.1
9	10	E	+1	15%	15.22%	59.9 (0.2)	60.0
5	10	B	+1	15%	15.22%	282.7 (0.9)	283.8
5	10	B	-1	15%	15.22%	279.5 (0.6)	279.2
10	20	E	+1	10%	10.35%	309.0 (0.9)	310.1
12	20	E	+1	10%	10.35%	253.9 (0.8)	254.5
14	20	E	+1	10%	10.35%	193.2 (0.6)	193.0
16	20	E	+1	10%	10.35%	129.3 (0.4)	129.2
18	20	E	+1	10%	10.35%	64.6 (0.2)	64.3
10	20	B	+1	10%	10.35%	347.8 (0.8)	351.3
10	20	B	-1	10%	10.35%	339.6 (0.9)	342.4

Table 1

In generating the Bermudan swaption prices in Table 1, the function $H(\cdot)$ was determined explicitly at all possible exercise dates. For medium- and long-dated options, this involves optimizing on a substantial number of variables (the 20-year swaption with 10-year lockout has,

[15] Notice in particular that convexity effects cause long-term swap rates tend to have lower effective volatility (over a finite time interval) in the BDT model than does the instantaneous short rate. This is the reason why a BDT model needs a 10.35% volatility to match the European swaption prices generated by a LM model with 10% volatility. The effect induces small but noticeable differences in the forward volatility structures of the two models

for instance, a total of 20 exercise dates) which again requires a relatively large number of pre-simulations n to produce a smooth exercise barrier. To improve speed, let us consider using the trick suggested in Section 2 and approximate the early exercise boundary by a piecewise linear function with q kink-points. In the table below, we list the simulated price of the 20-year Bermudan swaption with a 10-year lockout as a function of q as well as the number of pre-simulations n. The table contains the prices obtained by both strategies I-II in (9)-(10).

Bermudan Swaption Prices in 1-Factor LM Model

$d_k = 0.5$; $F_k(0) = 6\%$; $q = 6\%$; $j(x) = x$; $l_k(t) = 10\%$; $N = 50{,}000$

$T_s = 10$; $T_x = 19.5$; $T_e = 20$

		Strategy I		Strategy II	
		Price (SD)	Price (SD)	Price (SD)	Price (SD)
q	n	$f = +1$	$f = -1$	$f = +1$	$f = -1$
1	500	347.4 (0.8)	339.4 (0.9)	347.4 (0.8)	339.4 (0.9)
1	5,000	346.9 (0.8)	339.3 (0.9)	346.9 (0.8)	339.2 (0.9)
2	500	347.6 (0.8)	339.4 (0.9)	347.7 (0.8)	339.1 (0.9)
2	5,000	347.7 (0.8)	339.0 (0.9)	347.7 (0.8)	339.2 (0.9)
5	500	347.3 (0.8)	338.9 (0.9)	347.1 (0.8)	337.5 (0.9)
5	5,000	347.8 (0.8)	339.6 (0.9)	347.7 (0.8)	339.6 (0.9)
10	500	347.6 (0.8)	339.2 (0.9)	347.1 (0.8)	337.8 (1.0)
10	5,000	347.1 (0.8)	339.2 (0.9)	347.1 (0.8)	339.3 (0.9)

Table 2

As is evident from the table, the computed Bermudan prices are remarkably insensitive to both the number of kink-points and the number of simulation paths. For instance, using with just one kink-point (i.e. assuming that the boundary function $H(\cdot)$ is a perfectly straight line for 10 years) and $n = 500$ simulation paths, the resulting prices are statistically indistinguishable from the ones reported in Table 1. Similar results hold for the other Bermudan swaptions in Table 1. Table 2 also reveals that in our 1-factor setting, the more complicated strategy II adds no extra value to the simpler strategy I. Finally, notice that if q is sufficiently high (5 or above), using just $n = 500$ simulation paths in Strategy II will generate suboptimal results, particularly, it appears, for receiver swaptions. As mentioned earlier, if the number of simulated paths is too low relative to the numbers of parameters to be optimized on, noise will affect the estimation procedure too much, resulting in a choppy exercise boundary and sometimes suboptimal prices.

For the payer and receiver Bermudan swaptions in Table 2, the figures below graphs the boundaries for strategy I for a few different values of q. The figures also contain the exercise boundaries (parametrized in intrinsic swap value) implied in the 200x200 BDT lattice. As one would expect from the numbers in Table 2, the exercise boundary is almost linear; this, in fact, seems to hold quite generally for a large range of market conditions and contract specifications.

Payer Swaption Exercise Boundary (Strategy I) vs. Number of Kink-Points q

$d_k = 0.5$; $F_k(0) = 6\%$; $q = 6\%$; $j(x) = x$; $l_k(t) = 10\%$; $N = 50{,}000$; $n = 5{,}000$

$T_s = 10$; $T_x = 19.5$; $T_e = 20$

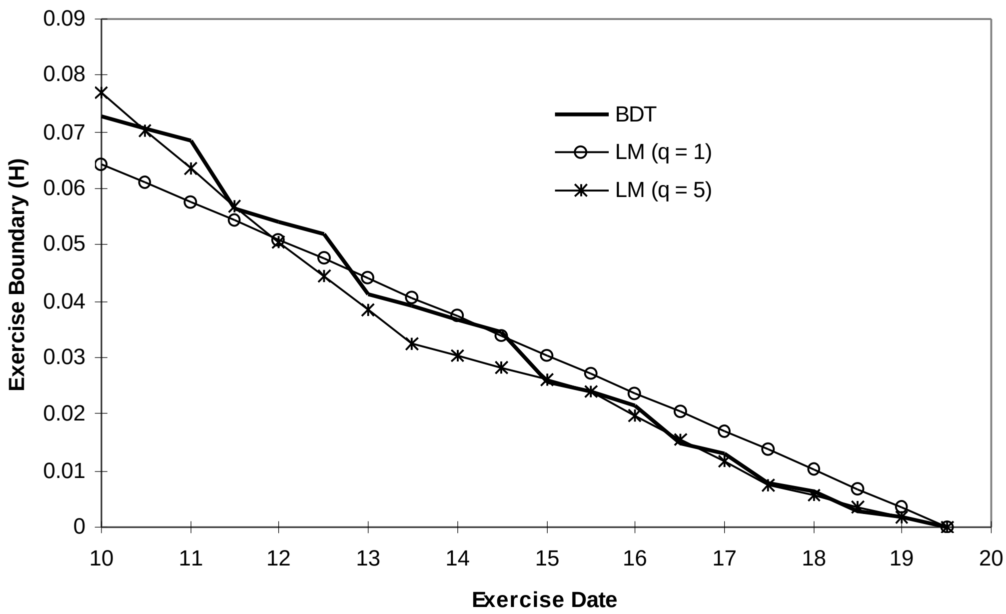

Figure 1

Receiver Swaption Exercise Boundary (Strategy I) vs. Number of Kink-Points q

$d_k = 0.5$; $F_k(0) = 6\%$; $q = 6\%$; $j(x) = x$; $l_k(t) = 10\%$; $N = 50{,}000$; $n = 5{,}000$

$T_s = 10$; $T_x = 19.5$; $T_e = 20$

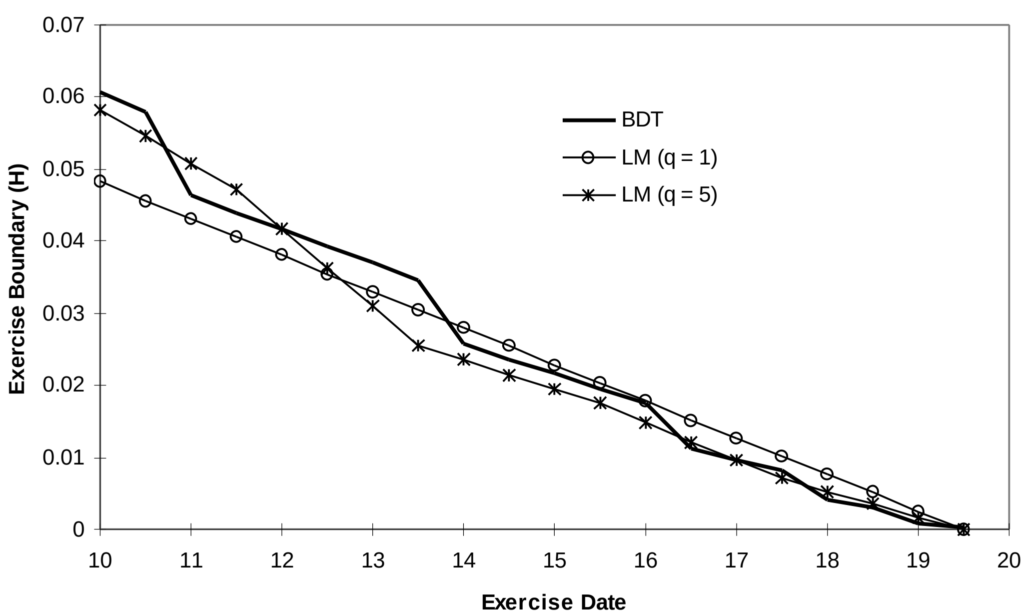

Figure 2

4.2 The sensitivity of prices with respect to the location of the exercise boundary.

As we saw in Table 2, the prices of Bermudan swaptions in our Monte Carlo algorithm appear quite insensitive to the exact location of the exercise barrier, a phenomenon that is well-known from the pricing of American equity options (see e.g. Ingersoll (1998) and Ju (1997)). To explore this further, consider now replacing the function $H(\cdot)$ (found in Step 3 of our algorithm) with $gH(\cdot)$, for some constant scaling factor g. In other words, after having located the exercise boundary we move it up or down by a certain multiplicative shift. In Table 3 below, we list the simulated price of the Bermudan swaptions in Table I, as a function of the scaling factor g. In calculating the optimal value of $H(\cdot)$ in the first place, we used $q = 5$ kink-points and $n = 5{,}000$ pre-simulations.

Sensitivity of Bermudan Swaption Prices (Strategy I) to Boundary Multiplier (g)

$d_k = 0.5$; $F_k(0) = 6\%$; $q = 6\%$; $j(x) = x$; $T_x = T_{e-1}$; $N = 50{,}000$; $n = 5{,}000$; $q = 5$

T_s	T_e	f	l_k	Price (SD) $g = 0.5$	Price (SD) $g = 0.75$	Price (SD) $g = 1.0$	Price (SD) $g = 1.25$	Price (SD) $g = 1.5$
1	4	+1	20%	155.2 (0.3)	157.4 (0.3)	157.7 (0.5)	155.6 (0.5)	152.6 (0.5)
1	4	-1	20%	152.8 (0.3)	155.8 (0.3)	156.6 (0.3)	155.5 (0.3)	152.5 (0.3)
2	5	+1	20%	185.8 (0.6)	187.5 (0.6)	187.9 (0.6)	186.8 (0.6)	184.9 (0.6)
2	5	-1	20%	183.9 (0.4)	185.9 (0.4)	186.6 (0.4)	186.3 (0.4)	184.5 (0.4)
5	10	+1	15%	279.9 (0.9)	282.2 (0.9)	282.6 (0.9)	280.6 (0.9)	277.4 (0.9)
5	10	-1	15%	276.3 (0.6)	278.7 (0.6)	279.4 (0.6)	278.1 (0.6)	274.9 (0.7)
10	20	+1	10%	342.3 (0.9)	346.4 (0.8)	347.6 (0.8)	346.1 (0.9)	341.9 (0.9)
10	20	-1	10%	335.0 (0.9)	338.5 (0.9)	339.6 (0.9)	338.6 (1.0)	335.0 (1.0)

Table 3

As Table 3 shows, the degree of accuracy needed in the estimation of $H(\cdot)$ is remarkably low: even if the location of the barrier is underestimated by a factor 2, the price of the Bermudan swaptions do not move by more than 3-4 basis points. The figure below graphs the price of the 20 year Bermudan payer swaption in Table 3 as a function of g; it is apparent that the function we are trying to optimize when searching for $H(\cdot)$ is quite flat around its extremum[16].

[16] Notice, that when $g = 0$ (the intercept of the graph with the price-axis), the option price does *not* equal the price of the corresponding European swaption. Rather, the price is that of a "flexi-swap" where the holder receives the first swap that is in-the-money. The price of this instrument is clearly higher than the price of the European swaption.

Bermudan Payer Swaption Price (Strategy I) vs. Boundary Multiplier g

$d_k = 0.5$; $F_k(0) = 6\%$; $q = 6\%$; $j(x) = x$; $l_k(t) = 10\%$; $N = 50{,}000$; $n = 5{,}000$; $q = 5$

$T_s = 10$; $T_x = 19.5$; $T_e = 20$

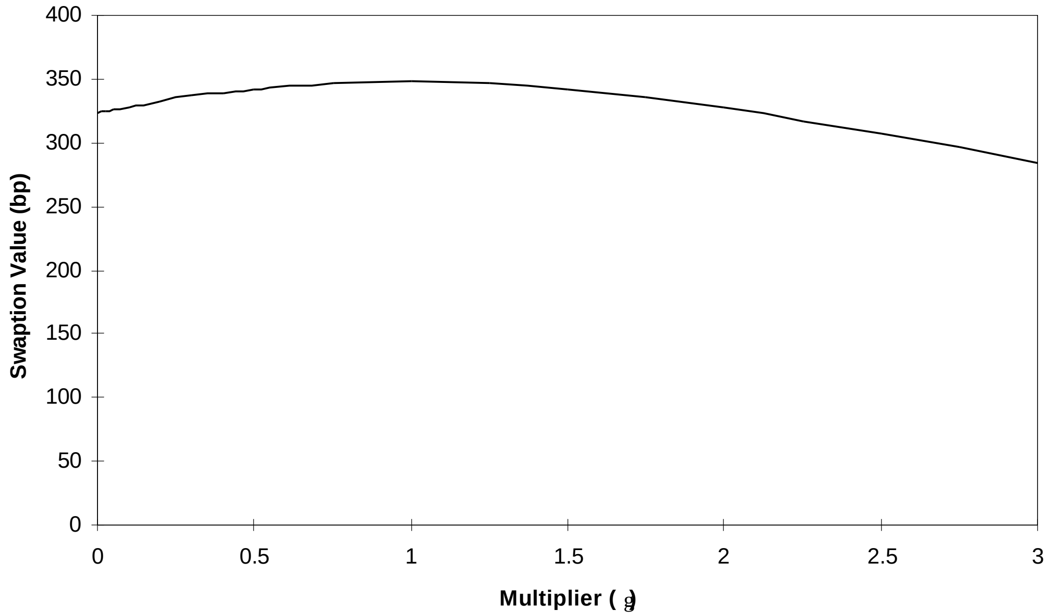

Figure 3

3.3 The effects of a skew.

To conclude this section, we will illustrate the effects of incorporating a volatility skew into the pricing of Bermudan swaptions. In particular, we will introduce a square-root CEV model with $j(x) = x^{1/2}$ in the driving SDE (2). As mentioned earlier, this will generate a downward-sloping volatility skew (see Footnote 3). In Table 4, we compare the Bermudan swaption prices obtained with this model against a log-normal specification. In an attempt to keep ATM prices roughly constant independent of the skew, we relate the CEV volatility l_k^{CEV} to the log-normal volatility (l_k^{LN}) as $l_k^{CEV} = l_k^{LN}\sqrt{F_k}$. As before, numbers in parentheses denote sample standard deviation.

Bermudan Swaption Prices (Strategy I) for different Coupon Rates (q)

$d_k = 0.5$; $F_k(0) = 6\%$; $j(x) = x^a$; $T_x = T_{e-1}$; $N = 50{,}000$; $n = 5{,}000$; $q = 5$

$T_s = 1$; $T_e = 4$; $l_k = 0.2 \cdot 0.06^{1-a}$

f	a	$q = 4\%$	$q = 5\%$	$q = 6\%$	$q = 7\%$	$q = 8\%$
+1	1	516.0 (0.2)	301.6 (0.4)	157.7 (0.5)	79.1 (0.4)	39.5 (0.3)
+1	1/2	518.4 (0.2)	306.2 (0.4)	159.2 (0.4)	76.1 (0.4)	33.7 (0.3)
-1	1	13.2 (0.1)	56.7 (0.2)	156.6 (0.3)	321.1 (0.3)	534.6 (0.2)
-1	1/2	17.3 (0.1)	60.1 (0.3)	155.7 (0.3)	316.6 (0.3)	530.3 (0.2)

$T_s = 2$; $T_e = 5$; $l_k = 0.2 \cdot 0.06^{1-a}$

f	a	$q = 4\%$	$q = 5\%$	$q = 6\%$	$q = 7\%$	$q = 8\%$
+1	1	498.7 (0.4)	315.5 (0.6)	187.9 (0.6)	108.8 (0.6)	60.5 (0.4)
+1	1/2	505.2 (0.4)	322.4 (0.5)	189.3 (0.5)	104.3 (0.5)	54.3 (0.4)
-1	1	23.3 (0.2)	80.04 (0.3)	186.6 (0.4)	341.6 (0.3)	532.8 (0.2)
-1	1/2	30.18 (0.2)	85.2 (0.4)	186.3 (0.4)	335.9 (0.4)	525.1 (0.3)

$T_s = 5$; $T_e = 10$; $l_k = 0.15 \cdot 0.06^{1-a}$

f	a	$q = 4\%$	$q = 5\%$	$q = 6\%$	$q = 7\%$	$q = 8\%$
+1	1	672.6 (0.6)	445.1 (0.8)	282.6 (0.9)	175.5 (0.8)	108.3 (0.7)
+1	1/2	686.0 (0.4)	455.9 (0.7)	284.1 (0.7)	167.3 (0.7)	93.6 (0.6)
-1	1	43.3 (0.3)	130.4 (0.5)	279.4 (0.6)	484.4 (0.6)	731.5 (0.4)
-1	1/2	56.3 (0.4)	139.7 (0.6)	279.4 (0.8)	475.1 (0.7)	717.1 (0.6)

$T_s = 10$; $T_e = 20$; $l_k = 0.1 \cdot 0.06^{1-a}$

f	a	$q = 4\%$	$q = 5\%$	$q = 6\%$	$q = 7\%$	$q = 8\%$
+1	1	865.4 (0.4)	562.8 (0.7)	347.6 (0.8)	208.0 (0.8)	123.2 (0.7)
+1	1/2	881.3 (0.3)	576.3 (0.6)	348.8 (0.7)	196.8 (0.7)	104.8 (0.6)
-1	1	45.9 (0.4)	150.4 (0.7)	339.6 (0.9)	606.1 (0.9)	930.0 (0.8)
-1	1/2	61.0 (0.5)	162.3 (0.9)	340.0 (1.1)	594.1 (1.2)	912.9 (1.1)

Table 4

Consistent with a downward sloping volatility skew, the distribution of forward rates under the square-root CEV process is skewed left of the log-normal distribution. As a consequence, in-the-money payer swaptions and out-of-the money receiver swaptions are priced higher in the square-root process than in the log-normal process. Conversely, prices of out-of-the-money payer swaptions and in-the-money receiver swaptions are higher in the log-normal model than in the square-root model. In the examples considered in Table 4, the effect of introducing a volatility skew is quite significant.

5. Two-factor models and comparisons with other approaches.

In this section we will compare our method against published results obtained with non-recombining trees and Markov chains. As part of these tests, we will investigate both one- and two-factor models. We have not included any results for models with more than two factors, mainly because no such results appear to be available in the literature. We do point out, however, that differences between two-factor models and models with a higher number of state variables are normally small; typically, a two-factor model will capture in excess of 95% of the overall yield curve variation.

5.1 Non-recombining trees.

Although quite a few papers about non-recombining trees can be found in the literature, very few contain concrete numerical results that can be reproduced independently. An exception[17] is a paper by Radhakrishnan (1998), which considers the pricing of certain Bermudan swaptions in the one- and two-factor log-normal HJM model. The log-normal HJM model is closely related to a log-normal LM model, but uses continuous, rather than discrete, compounding conventions in the computation of forward rates. The two models can normally be brought closely in line by a slight upward scaling of the HJM volatilities. We will consider two scenarios (here denoted A and B) from Radhakrishnan (1998), the first involving one Brownian motion, the second two. Both scenarios involve a semi-annual forward curve of $F_k(0) = 5.063\%$ and log-normal dynamics ($j(x) = x$). Appendix A gives more details about the volatility structures in the two scenarios and also lists the scaling factors (found empirically by matching the European option prices listed in Table 6B of Radhakrishnan (1998)) that were employed to convert the HJM volatilities to LM form.

Below, we list European and Bermudan receiver swaption prices for scenarios A and B. Tables 5a and 5b below contain both the prices computed by Radhakrishnan's non-recombining trees[18] (taken from his Table 6B), as well as the prices obtained by our Monte Carlo ("MC") strategies I and II (in (9) and (10)). Notice that the swaptions in tables 5a-b are non-standard in the sense that there is no lock-out period, $T_s = 0$, and that the last exercise date $T_x = 3$ is several periods before the terminal swap maturity[19].

---

[17] Another exception is a paper by Carr and Yang (1998), also dealing with the HJM model. This paper, however, does not discuss options on swaps and instead focuses on short-term options on zero-coupon bonds.

[18] Radhakrishnan uses a two-branch tree with 20 time steps for Scenario A, and a three-branch tree with 14 time step in Scenario B. The magnitude of the discretization errors on Radhakrishnan's are, of course, difficult to estimate, but the accuracy is probably adequate for the relatively short (3-year) exercise horizons of the options considered here.

[19] While Radhakrishnan describes this type of structure as a "regular Bermudan swaption", this is not the convention used in the market (see our discussion at the end of Section 2).

Bermudan and European Swaption Prices (in Basis Points) for Volatility Scenario A
$d_k = 0.5$; $F_k(0) = 5.063\%$; $j(x) = x$; $N = 50{,}000$; $n = 5{,}000$; $q = 4$
$T_s = 0$; $T_x = 3$; $f = -1$

		European Price (SD)		Bermudan Price (SD)		
T_e	q	20-Step Tree	MC	20-Step Tree	MC (Strat. I)	MC(Strat. II)
5	4%	34.7	34.8 (0.2)	47.9	49.3 (0.2)	49.3 (0.2)
5	5%	103.4	103.9 (0.3)	176.0	175.6 (0.3)	175.6 (0.3)
5	6%	210.5	210.2 (0.2)	424.1	421.2 (0.4)	420.9 (0.4)
8	4%	79.4	79.1 (0.5)	90.0	90.6 (0.5)	91.0 (0.5)
8	5%	237.2	239.1 (0.7)	298.0	299.2 (0.6)	298.6 (0.6)
8	6%	486.9	486.3 (0.5)	676.2	674.7 (0.8)	675.2 (0.8)

The maturity of all European swaptions above is 3 years.

Table 5a

Bermudan and European Swaption Prices (in Basis Points) for Volatility Scenario B
$d_k = 0.5$; $F_k(0) = 5.063\%$; $j(x) = x$; $N = 50{,}000$; $n = 5{,}000$; $q = 4$
$T_s = 0$; $T_x = 3$; $f = -1$

		European Price (SD)		Bermudan Price (SD)		
T_e	q	14-Step Tree	MC	14-Step Tree	MC (Strat. I)	MC(Strat. II)
5	4%	34.6	35.0 (0.2)	47.3	47.9 (0.2)	48.1 (0.2)
5	5%	103.1	104.0 (0.3)	170.5	171.4 (0.3)	172.1 (0.3)
5	6%	208.6	209.9 (0.2)	412.8	415.3 (0.4)	415.6 (0.4)
8	4%	78.6	75.3 (0.5)	87.8	82.6 (0.4)	83.6 (0.4)
8	5%	233.8	233.4 (0.6)	284.9	282.1 (0.5)	284.0 (0.5)
8	6%	477.0	480.8 (0.5)	653.7	654.0 (0.7)	655.1 (0.8)

The maturity of all European swaptions above is 3 years.

Table 5b

Overall, the results in tables 5a-b are good, with the Monte Carlo prices of the Bermudan swaptions typically differing only a few basis points from the prices generated by the trees. Given that the European option prices are not matched exactly either (particularly for the two-factor Scenario B), we point out that some, if not most, of the observable price differences are likely due to differences between the dynamics of the HJM and "fitted" LM models, rather than to differences in the numerical techniques used. As expected, Monte Carlo strategies I and II give virtually identical results for the one-factor Scenario A. For the two-factor Scenario B, Strategy II results in a slight improvement (1-2 basis points) over Strategy I. Despite its simplicity, Strategy I thus appears to hold up quite well, even in the case of multi-factor models.

5.2 Comparison with the Markov chain method.

In this section we will compare our method with the Markov Chain technique of Carr and Yang (1997). For our examples, we set the quarterly compounded yield curve to 10% flat ($d_k = 0.25$, $F_k(0) = 10\%$ for all k), assume log-normal evolution of forward rates ($j(x) = x$), and consider the following two volatility scenarios:

Scenario C (1-Factor Model): $l_k(t) = 0.2$ for all k and $t \le T_k$

Scenario D (2-Factor Model): $l_k(t) = \left(0.15,\ 0.15 - \sqrt{0.009(T_k - t)}\right)^T$, $t \le T_k$

The tables below list compare European and Bermudan payer swaption prices listed in Carr and Yang (1997) by those obtained by Monte Carlo simulation ("MC") of our exercise strategies I and II. While Carr and Yang report prices for a variety of (non-standard) structures, we here focus on regular Bermudan swaptions[20] with allowed exercise at all coupon dates after an initial lock-out period. As before, numbers in parentheses denote sample standard deviations[21].

Bermudan and European Swaption Prices (in Basis Points) for Volatility Scenario C
$d_k = 0.25$; $F_k(0) = 10\%$; $j(x) = x$; $T_x = T_{e-1}$; $f = +1$; $N = 50{,}000$; $n = 5{,}000$; $q = 4$

			European Price (SD)		Bermudan Price (SD)		
T_s	T_e	q	Carr/Yang	MC	Carr/Yang	MC (Strat. I)	MC (Strat. II)
0.25	1.25	8%	183.9 (0.0)	183.8 (0.0)	184.7 (0.0)	184.6 (0.1)	184.6 (0.1)
0.25	1.25	10%	36.6 (0.0)	36.5 (0.1)	49.2 (0.0)	49.1 (0.1)	49.1 (0.1)
0.25	1.25	12%	1.3 (0.0)	1.3 (0.0)	8.7 (0.0)	8.9 (0.1)	8.9 (0.1)
1	3	8%	344.3 (0.0)	343.9 (0.3)	355.9 (0.1)	355.6 (0.4)	355.6 (0.4)
1	3	10%	129.5 (0.2)	129.7 (0.5)	157.5 (0.1)	157.8 (0.5)	156.6 (0.4)
1	3	12%	34.8 (0.1)	35.1 (0.3)	61.8 (0.2)	61.8 (0.4)	61.8 (0.4)
1	6	8%	748.4 (0.1)	747.6 (0.6)	811.2 (0.5)	807.2 (0.9)	805.6 (0.8)
1	6	10%	281.4 (0.4)	281.7 (1.0)	419.9 (0.4)	417.8 (0.9)	417.7 (0.9)
1	6	12%	75.5 (0.1)	76.1 (0.7)	215.8 (0.5)	212.7 (0.9)	212.7 (0.9)
1	11	8%	1205.0 (0.2)	1203.4 (0.7)	1395.2 (0.8)	1381.6 (1.6)	1381.6 (1.6)
1	11	10%	452.6 (0.6)	452.9 (1.5)	830.1 (0.7)	812.9 (1.4)	812.9 (1.4)
1	11	12%	121.2 (0.2)	122.1 (1.0)	511.9 (0.9)	495.8 (1.5)	495.1 (1.5)
3	6	8%	473.8 (0.2)	472.6 (0.8)	495.5 (0.3)	493.7 (0.8)	491.2 (0.8)
3	6	10%	262.5 (0.3)	262.4 (1.0)	295.1 (0.3)	294.6 (0.9)	293.7 (0.9)
3	6	12%	136.3 (0.2)	136.3 (0.9)	171.1 (0.3)	170.3 (0.8)	170.7 (0.8)

Table 6a

[20] Carr and Yang use the term "fixed-tail" swaption for this structure.

[21] The numbers of Carr and Yang were produced by 100 batches of 100,000 Markov chain paths

Bermudan and European Swaption Prices (in Basis Points) for Volatility Scenario D
$d_k = 0.25$; $F_k(0) = 10\%$; $j(x) = x$; $T_x = T_{e-1}$; $f = +1$; $N = 50{,}000$; $n = 5{,}000$; $q = 4$

			European Price (SD)		Bermudan Price (SD)		
T_s	T_e	q	Carr/Yang	MC	Carr/Yang	MC (Strat. I)	MC (Strat. II)
0.25	1.25	8%	183.7 (0.0)	183.6 (0.0)	184.2 (0.0)	184.0 (0.0)	184.0 (0.0)
0.25	1.25	10%	31.6 (0.0)	31.6 (0.1)	43.3 (0.0)	43.3 (0.1)	43.4 (0.1)
0.25	1.25	12%	0.5 (0.0)	0.6 (0.0)	5.5 (0.0)	5.6 (0.1)	5.6 (0.1)
1	3	8%	333.1 (0.0)	332.2 (0.2)	341.0 (0.1)	339.7 (0.2)	339.8 (0.2)
1	3	10%	101.5 (0.1)	101.1 (0.4)	127.0 (0.1)	125.8 (0.3)	125.9 (0.3)
1	3	12%	16.5 (0.1)	16.7 (0.2)	37.1 (0.1)	36.9 (0.2)	36.8 (0.2)
1	6	8%	721.4 (0.1)	719.8 (0.3)	750.0 (0.3)	750.2 (0.6)	749.6 (0.6)
1	6	10%	211.9 (0.3)	211.3 (0.8)	316.7 (0.4)	317.0 (0.7)	315.9 (0.7)
1	6	12%	31.2 (0.1)	31.6 (0.4)	127.9 (0.3)	127.7 (0.6)	128.0 (0.6)
1	11	8%	1165.9 (0.2)	1163.7 (0.4)	1221.0 (0.8)	1247.3 (1.2)	1250.9 (1.2)
1	11	10%	353.5 (0.4)	352.5 (1.2)	602.2 (0.9)	620.8 (1.1)	627.1 (1.1)
1	11	12%	55.8 (0.2)	56.8 (0.6)	323.4 (0.8)	327.1 (1.2)	331.8 (1.1)
3	6	8%	431.8 (0.1)	429.8 (0.5)	447.2 (0.2)	444.7 (0.6)	444.4 (0.6)
3	6	10%	200.7 (0.2)	199.9 (0.8)	228.2 (0.2)	226.9 (0.7)	227.2 (0.7)
3	6	12%	79.6 (0.1)	79.5 (0.6)	107.7 (0.2)	107.1 (0.6)	107.1 (0.6)

Table 6b

Generally speaking, the numbers in Carr and Yang (1997) are quite close to those generated by our Monte Carlo methods -- typically the Bermudan swaption prices are within a few basis points. An exception, however, occurs for the 1-year option on 10-year swaps. For the one-factor Scenario C, Carr and Yang's Bermudan prices are significantly above (up to 17 basis points) the Monte Carlo results; for the two-factor scenario, their results are significantly below (up to 26 basis points) those obtained by Monte Carlo simulation. As the numbers generated by our Monte Carlo simulation are a low estimate of the true price, it appears reasonable to conclude that Carr and Yang's Markov chain method loses some accuracy for long-tenor swaptions in a two-factor setting. In the case of the one-factor model, it is less clear whether there is a problem with the Markov chain approach, or whether our low estimate is just far below the right price. Given that our previous results from sections 4 and 5.1 indicate that our method is quite accurate for one-factor models we tend to believe that the price differences are due to a loss of accuracy in the Markov chain approach, rather than in our method. In general, the maturities and swap tenors used by Carr and Yang are relatively modest; it would be interesting to compare the two approaches for Bermudan swaptions with maturities and swap tenors in excess of, say, 5-10 years.

As expected, for the one-factor Scenario C our exercise strategy II offers no benefits whatsoever over the simpler strategy I. With two factors (Scenario D), the improvements are also modest or non-existing, except for the 1-year option on 10-year swaps. Here, Strategy II results in a pickup of roughly 5 basis points relative to Strategy I. Intuitively, for the correlation effects introduced by two-factor model to matter, the exercise period must be quite long; otherwise, even a two-factor model would imply near-perfect correlation of the different swaps the option holder can exercise into.

5.3. More numerical results.

To conclude this section we will list computed prices of a few medium- to long-dated ATM Bermudan swaptions. All numbers in the table below were generated using a semi-annual forward curve equal to 6% and a two-factor log-normal ($\varphi(x) = x$) volatility structure given by

Scenario E (2-factor model): $\lambda_k(t) = \left(0.10,\ 0.10 - \sqrt{0.002(T_k - t)}\right)^{\mathrm{T}}$

Bermudan and European Swaption Prices (in Basis Points) for Volatility Scenario E
$\delta_k = 0.5$; $F_k(0) = 6\%$; $q = 6\%$; $\varphi(x) = x$; $T_x = T_{e-1}$; $N = 50{,}000$; $n = 10{,}000$; $q = 5$

T_s	T_e	(B)ermudan (E)uropean	ϕ	Strategy I Price (SD)	Strategy II Price (SD)
3	8	E	+1	151.0 (0.6)	151.0 (0.6)
3	8	B	+1	182.8 (0.5)	183.1 (0.5)
3	8	B	-1	181.0 (0.4)	181.1 (0.4)
3	13	E	+1	259.6 (0.9)	259.6 (0.9)
3	13	B	+1	350.5 (0.8)	352.1 (0.8)
3	13	B	-1	343.2 (0.7)	343.5 (0.7)
5	10	E	+1	170.7 (0.6)	170.7 (0.6)
5	10	B	+1	194.4 (0.6)	194.4 (0.6)
5	10	B	-1	192.6 (0.5)	192.7 (0.5)
5	15	E	+1	299.0 (1.0)	299.0 (1.0)
5	15	B	+1	367.3 (0.9)	368.9 (0.9)
5	15	B	-1	360.0 (0.8)	360.5 (0.8)
10	15	E	+1	184.3 (0.7)	184.3 (0.7)
10	15	B	+1	197.2 (0.6)	197.4 (0.6)
10	15	B	-1	196.2 (0.5)	196.2 (0.5)
10	20	E	+1	331.5 (1.0)	331.5 (1.0)
10	20	B	+1	369.2 (1.0)	370.0 (1.0)
10	20	B	-1	362.1 (1.0)	361.5 (1.0)

Table 7

6. Conclusions.

In this paper, we have described a fast and robust technique to compute the prices of Bermudan swaptions by Monte Carlo simulation in the Libor Market model. The method is based on an explicit description of the exercise boundary and is simple to implement, yet gives surprisingly good results. We discussed two specific exercise strategies and demonstrated that option prices are remarkably insensitive to the accuracy with which the exercise boundary is estimated. Moreover, it appeared that a simple strategy based on intrinsic value is adequate for most options, particularly if the number of driving Brownian Motions is low. While the prices generated by the method are biased low, experimental comparisons with bias-free tree and lattice methods suggest that the bias is very small and certainly within any reasonable bid-offer spread. Our results were also comparable to those of Carr and Yang (1997), although certain discrepancies were observed for options on long-dated swaps. While the lack of feasible bias-free pricing methods makes it difficult to establish the source of these discrepancies, it appears that the Markov chain method of Carr and Yang (1997) might lose some accuracy for long-dated swaps.

We should point out that while this paper has exclusively dealt with Bermudan swaptions, the techniques presented can easily be extended to other, more complicated, Bermuda-style interest rate derivatives. For instance, we have successfully applied the proposed technique to the pricing of callable reverse floaters and Bermuda-style options on caps.

As a final comment, let us make it clear that the simple technique presented in this paper is largely motivated by practical considerations and the need to "get something done". Much work remains, particularly on the establishment of a practical technique to determine tight upper bounds on Bermudan derivatives prices. Hopefully the technique and, in particular, the numerical results in this paper will prove useful for future research.

References.

Andersen, L. and J. Andreasen (1998), "Volatility Skews and Extensions of the Libor Market Model," Working Paper, General Re Financial Products.

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

Black, F., E. Derman, and W. Toy (1990), "A One-Factor Model of Interest Rates and Its Applications to Treasury Bond Options, " *Financial Analysts Journal*, January-February, 33-39.

Black, F. and P. Karasinski (1991), Bond and Option Pricing when Short Rates are Lognormal," *Financial Analysts Journal*, July-August, 52-59.

Boyle, P., M. Broadie, and P. Glasserman (1997). "Simulation Methods for Security Pricing," *Journal of Economic Dynamics and Control*, 21, 1267-1321.

Brace, A., M. Gatarek, and M. Musiela (1997). "The Market Model of Interest Rate Dynamics," *Mathematical Finance*, 7, 127-155.

Broadie, M. and P. Glasserman (1997a), "Pricing American-Style Securities Using Simulation," *Journal of Economic Dynamics and Control*, 21, 8-9, 1323-1352..

Broadie, M. and P. Glasserman (1997b), "A Stochastic Mesh Method for Pricing High-Dimensional American Options," Working Paper, Columbia University.

Carr P. and G. Yang (1997), "Simulating Bermudan Interest Rate Derivatives," Working Paper, Morgan Stanley / NumeriX.

Carr P. and G. Yang (1998), "Simulating American Bond Options in an HJM Framework," Working Paper, Morgan Stanley / NumeriX.

Carverhill, A. (1995), "A Note on the Models of Hull and White for Pricing Options on the Term Structure: Response," *Journal of Fixed Income*, 5, 2, 89-96.

Clewlow, L. and C. Strickland (1998), "Pricing Interest Rate Exotics by Monte Carlo Simulation," Chapter 17, *Monte Carlo: Methodologies and Applications for Pricing and Risk Management*, RISK Books, London.

Gatarek, D. (1996), "Pricing of American Swaptions in the Approximate Lognormal Model," Working Paper, University of New South Wales.

Grant, D., G. Vora, and D. Weeks (1997), "Path-Dependent Options: Extending the Monte Carlo Simulation Approach," *Management Science*, 43, 11, 1589-1602.

Heath, D., R. Jarrow, and A. Morton (1990). "Contingent Claim Valuation with a Random Evolution of Interest Rates," *Review of Futures Markets*, 9, 1, 54-76.

Heath, D., R. Jarrow, and A. Morton (1992). "Bond Pricing and the Term Structure of Interest Rates: A new Methodology for Contingent Claims Evaluation," *Econometrica*, 60, 77-105.

Hull, J. and A. White (1990), "Pricing Interest Rate Derivative Securities," *Review of Financial Studies*, 3, 4, 573-592.

Hull, J. and A. White (1995), "A Note on the Models of Hull and White for Pricing Options on the Term Structure: Response," *Journal of Fixed Income*, 5, 2, 97-102.

Ingersoll, J. (1998), "Approximating American Options and other Financial Contracts using Barrier Derivatives," *Journal of Computational Finance*, 2, 1, 85-112.

Jamshidian, F. (1997). "Libor and Swap Market Models and Measures," *Finance and Stochastics*, 1, 293-330.

Ju, N. (1997), "Pricing an American Option by Approximating its early Exercise Boundary as a Piece-wise Exponential Function," Working Paper, Haas School of Business, University of California at Berkeley.

Li, B. (1996), "Latest Advances in Monte Carlo Methods for Pricing and Hedging American Style Options," RISK Conference Notes, Merrill Lynch.

Longstaff, F. and E. Schwarz, "Valuing American Options by Simulation: A Simple Least-Squares Approach, " Working Paper, Anderson School of Business, UCLA.

Miltersen, K., K. Sandmann, and D. Sondermann (1997). "Closed-Form Solutions for Term Structure Derivatives with Lognormal Interest Rates," *Journal of Finance*, 409-430.

Pedersen, M. (1999), "Bermudan Swaptions in the Libor Market Model," Working Paper, SimCorp A/S.

Press, W., S. Teukolsky, W. Vetterling, and B. Flannery (1992). *Numerical Recipes in C*, Cambridge University Press.

Radhakrishnan, A. (1998), "Does Correlation Matter in Pricing Caps and Swaptions," Working Paper, Stern School of Business, New York University.

Raymar, S. and M. Zwecher (1997), "Monte Carlo Estimation of American Call Options on the Maximum of Several Stocks," *Journal of Derivatives*, 5, 1 (1997).

Ritchken, P. and L. Sankarasubramanian (1995), "Volatility Structures of Forward Rates and the Dynamics of the Term Structure," *Mathematical Finance*, 5, 55-72.

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

Appendix A

The tables below list the HJM volatilities of the scenarios A and B (from Radhakrishnan (1998), Table 5)[22]. The scaling factors for conversion of HJM volatilities to LM form were found empirically by matching the European swaption prices listed in Radhakrishnan (1998), Table 6B. The resulting scaling factors for scenarios A and B were 1.02 and 1.035, respectively.

Volatility Scenario A (1-factor model)

$l_k(t) = 1.02 \cdot s^{HJM}(T_k - t)$, $s^{HJM}(\cdot)$ in table below

$T_k - t$	$s^{HJM}(T_k - t)$
0	0.244
1	0.201
2	0.191
3	0.185
4	0.187
5	0.190
6	0.189
7	0.189
8	0.190

Volatility Scenario B (2-factor model)

$l_k(t) = 1.035 \cdot \left(s_1^{HJM}(T_k - t),\ s_2^{HJM}(T_k - t) \right)^T$; $s_1^{HJM}(\cdot)$ and $s_2^{HJM}(\cdot)$ in table below

$T_k - t$	$s_1^{HJM}(T_k - t)$	$s_2^{HJM}(T_k - t)$
0	0.207	-0.130
1	0.199	-0.026
2	0.189	0.026
3	0.173	0.065
4	0.156	0.104
5	0.138	0.130
6	0.123	0.143
7	0.113	0.152
8	0.104	0.159

The forward rate correlation matrix consistent with Scenario B can be found in Table 5 in Radhakrishnan (1998).

[22] We assume that entries between cells in the two tables below are to be computed by linear interpolation.