Volatility Skews and Extensions of the Libor Market Model[1]

Leif Andersen and Jesper Andreasen

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

First Version: August 12, 1997
This Version: August 27, 1998

Abstract

This paper considers extensions of the *Libor market model* (Brace *et al* (1997), Jamshidian (1997), Miltersen *et al* (1997)) to markets with volatility skews in observable option prices. We expand the family of forward rate processes to include diffusions with non-linear forward rate dependence and discuss efficient techniques for calibration to quoted prices of caps and swaptions. Special emphasis is put on generalized CEV processes for which closed-form expressions for cap and swaption prices are derived. We also discuss modifications of the CEV process which exhibit more appealing growth and boundary characteristics. The proposed models are investigated numerically through Crank-Nicholson finite difference schemes and Monte Carlo simulations.

1. Introduction

In a significant new line of research, the recent papers by Brace *et al* (1997), Jamshidian (1997), and Miltersen *et al* (1997) introduce a novel approach to arbitrage-free term structure modeling. Rather than working with the continuously compounded instantaneous forward rates as in Heath *et al* (1992), or the continuously compounded spot interest rates as in Vasicek (1977) and Cox *et al* (1985), these papers take discretely compounded (Libor) forward rates as the model primitives. Unlike continuously compounded forward rates, log-normally diffused discrete forward rates turn out to be non-explosive and, significantly, allow for pricing of Libor caplets by the "market convention" Black (1976) formula. The log-normal models advocated by Brace *et al* (1997), Jamshidian (1997), and Miltersen *et al* (1997) are therefore often termed *Libor market models*.

[1] The authors wish to thank Steven Shreve, Paul Glasserman, Wes Petersen, and Jakob Sidenius for insights and discussions.

While the Libor market models do not allow for usage of the Black (1976) formula in the pricing of swaptions, Brace *et al* (1997) derive good closed-form approximations for swaption prices under the log-normal market model assumptions. Availability of closed-form pricing formulas for both caps and swaptions enables efficient calibration of the model to market prices, a key feature of the model in terms of its usefulness in practical applications.

The basic premise of the Libor market model -- log-normally distributed Libor rates -- is, however, increasingly being violated in many important cap and swaption markets. In particular, implied Black (1976) volatilities of caplet and swaption prices often tend to be decreasing functions of the strike and coupon, respectively, indicating a fat left tail of the empirical forward rate distributions relative to log-normality. This so-called *volatility skew* is currently most pronounced in the Japanese Libor market, but also exists in US and German markets, among others. The presence of the volatility skew motivates the formulation of models where the diffusion coefficients of the discrete forward rates are non-linear functions of the rates themselves. In this paper we describe a general class of such models, which we will term *extended market models*. The models focused on here are characterized by a forward rate diffusion term that is *separable*, in the sense that it can be described as a product of a general time- and maturity-dependent function and a time-homogeneous non-linear function of the forward rate.

The separable form of the diffusion coefficient is shown to be tractable and allows for quick calibration to caplets by numerical solution of one-dimensional forward or backward partial differential equations (PDEs). For this we suggest an efficient numerical routine based on a deterministic time-change and the Crank-Nicolson finite-difference scheme. Alternatively, for the case where the forward-dependence of the diffusion term can be described by a power function, also termed the *CEV* (Constant Elasticity of Variance) model, we derive closed-form solutions for caplet prices. These results essentially extend the analysis of Schroder (1989) to the time-inhomogeneous case.

As we will show, the CEV model is about as tractable as the log-normal market model but can provide a much closer fit to observed caplet prices. To motivate our studies of the CEV model, below we show implied Black (1976) volatilities of CEV model caplet prices as functions of strike plotted against bid and ask implied caplet volatilities from the Japanese Libor market (provided by the GRFP interest rate option desk, May 1998). We have included prices for 2- and 9-year caplets; the CEV power (to be defined later) of the volatility coefficient is set to 0.6 for both maturities.

Market and CEV caplet prices in Japanese Libor market, May 1998

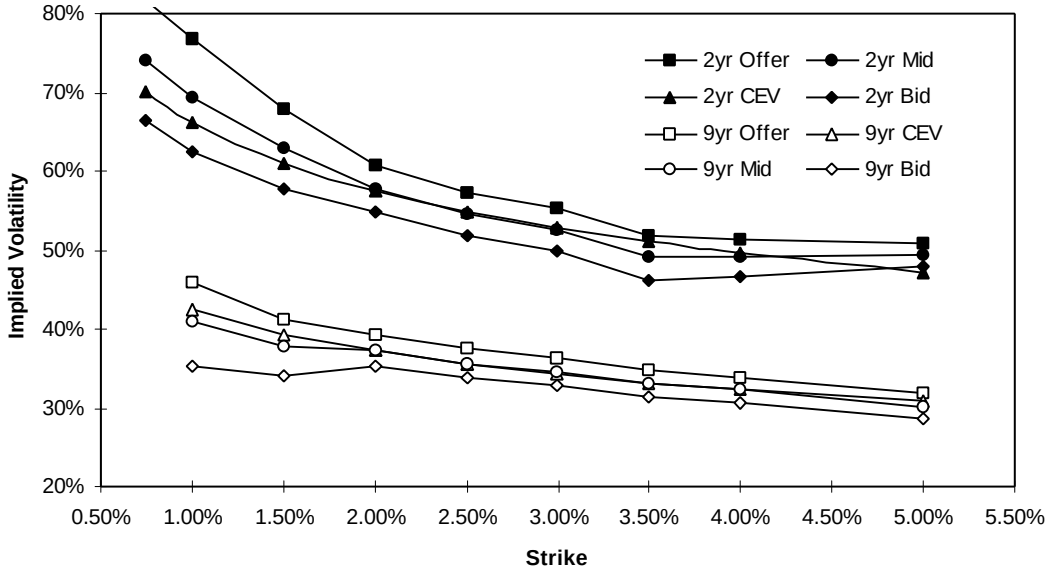

Figure 1

Though closed-form caplet prices and a good market fit makes the CEV model attractive it also exhibits certain technical irregularities. These can be circumvented, however, by the introduction of a 'regularized' version of the CEV process, here named the *LCEV* (*L*imited CEV) model. We show that the CEV closed-form caplet prices can be seen as a limiting case of those produced by the LCEV model. By numerical examples we illustrate that the CEV formulas can be used as very accurate approximations of caplet prices under the LCEV process.

For swaptions, the market model is less tractable than is the case for caps and floors. By making certain simplifying assumptions, however, we demonstrate that swaptions can be treated in exactly the same way as caplets. In particular, we are able to construct highly accurate closed-form approximations for swaption prices in the CEV market model. Our analysis is based on the concept of forward swap measures (see Jamshidian (1997)) which simplifies the development of closed-form approximations significantly compared to the approach taken in Brace *et al* (1997).

In the final part of the paper, we consider schemes to implement the proposed framework in a Monte Carlo setting. Monte Carlo simulations are then used to examine some of our results through numerical examples. Particular emphasis is put on tests of the swaption approximations and on quantifying discretization biases.

The rest of this paper is organized as follows. In Section 2 we provide notation and introduce the probability measures and stochastic processes necessary for later work. In Section 3, we narrow the discussion to the class of "separable" forward rate processes. After proving certain existence and uniqueness results, we describe a technique of deterministic time-change that

proves useful for this class of models. The section also introduces the CEV process and derives its transition density. In Section 4 we consider the PDEs for pricing of caplets and derive closed-form formulas for caplet prices in the CEV model. We also introduce the LCEV model and consider the convergence of the LCEV to the CEV model. Section 5 discusses the pricing of swaptions using closed-form approximations, and Section 6 is devoted to Monte-Carlo implementation of the extended market models and various numerical tests. Finally, Section 7 contains our conclusions. For clarity, all significant proofs are deferred to an appendix.

2. Basic Setup

Consider an increasing maturity structure $0 = T_0 < T_1 < ... < T_{K+1}$ and define a right-continuous[2] mapping function $n(t)$ by

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

While we do not put any restrictions on the maturity structure other than it being increasing, in practice we would often use a nearly equidistant spacing between points (say 3 or 6 calendar months) to match conventions used in swap and futures markets. With $P(t,T)$ denoting the time *t* price of a zero-coupon bond maturing at time *T*, we define discrete 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.

The discrete forward rates constitute our model primitives and collectively determine the state and evolution of interest rates. To state our assumptions about the stochastic processes driving the forward rates, we first fix our probability measure to be the T_{k+1} forward measure $\mathbf{Q}^{k+1}$, i.e. the equivalent probability measure induced by using the T_{k+1}-maturity zero-coupon

[2] While some authors define $n(t)$ to be left-continuous, we find our definition more convenient, particularly for discrete-time numerical work. In particular, our definition ensures that $n(t)$ does not jump when we move forward from a date that coincides with a point in the maturity structure.

bond as numeraire. We will assume that $\mathbf{Q}^{k+1}$ exist and is unique for all *k*. Absence of arbitrage then implies that $P(t,T_k)/P(t,T_{k+1})$, and thus $F_k(t)$, are martingales. Assuming that the forward rate dynamics are governed by a vector Brownian motion, we specify the forward rate dynamics as an Ito process:

$$dF_k(t) = s_k(t)^{\mathrm{T}} dW_{k+1}(t), \tag{1}$$

where $W_{k+1}(t)$ is an *m*-dimensional Brownian motion under $\mathbf{Q}^{k+1}$ and $s_k(t)$ is an *m*-dimensional adapted volatility function satisfying the usual integrability conditions. The forward measures corresponding to the different times of the maturity structure can be shown (Jamshidian (1997)) to be related iteratively through:

$$dW_{k+1}(t) = dW_k(t) + \frac{d_k s_k(t)}{1+ d_k F_k(t)} dt .$$

Defining the convenient *spot measure* $\mathbf{Q}$ as the equivalent measure under which $W(t) = W_{n(t)}(t)$ is a Brownian motion, we have from (1)

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

Notice, that the numeraire of the spot measure should not be interpreted as $P(t,n(t))$ but rather as a "rolling" zero-coupon bond (Jamshidian (1997)) with time *t* price of

$$B(t) = P(t,T_{n(t)}) \prod_{j=0}^{n(t)-1} P(T_j,T_{j+1})^{-1} = P(t,T_{n(t)}) \prod_{j=0}^{n(t)-1} (1+ d_j F_j(T_j)) . \tag{3}$$

While not used in this paper, we point out that the measure $\mathbf{Q}^{K+1}$ also receives special attention in the literature and is known as the *terminal measure*.

3. The Libor Market Model and its Extensions

The framework set up in Section 2 so far is quite broad and, through the choice of volatility functions $s_k(t)$, allows for a variety of models for interest rate evolution. For example, in the popular *Libor market model* (Brace *et al* (1997), Jamshidian (1997), Miltersen *et al* (1997)), the volatility functions are of the form

$$s_k(t) = F_k(t)\, l_k(t),$$

where $l_k(t)$ is a bounded *m*-dimensional deterministic function. The resulting log-normal distribution of $F_k(t)$ under the T_{k+1} forward measure justifies the usage of the market standard Black (1976) formula for interest rate caps and floors. While the function $l_k(t)$ allows for calibration to a term-structure of implied caplet volatilities, the market model formulation is not rich enough to capture the often observed dependence of implied interest rate option volatilities on strike (the volatility skew). This motivates formulation of models in which the volatility functions are specified as

$$s_k(t) = j(F_k(t))\, l_k(t). \tag{4}$$

where again $l_k(t)$ is a bounded vector valued deterministic function, and $j : [0,\infty) \to [0,\infty)$ is a possibly non-linear function.

Theorem 1.
Suppose that $j(0) = 0$ *and* $F_k(0) \geq 0,\ 0 \leq k \leq K$. *If in (4)* j *satisfies*

a) Local Lipschitz Continuity:

$$\forall n > 0,\ \exists C_n > 0 \ s.t. \ if \ \ 0 \leq x < n \ and \ \ 0 \leq y < n, \ then \ \ |j(x) - j(y)| \leq C_n |x - y|,$$

b) Linear Growth:

$$\exists C > 0 \ s.t. \ \ j(x)^2 \leq C(1 + x^2),\ \forall x > 0,$$

then non-explosive, pathwise unique solutions of the no-arbitrage SDEs for $F_k(t)$, $n(t) \leq k \leq K$, *exist under all measures* $\mathbf{Q}^i$, $n(t) \leq i \leq K+1$. *If* $F_k(0) > 0$, *the solutions are positive for all t >* 0.

Proof:
In Appendix A.♠

The local Lipschitz condition in Theorem 1 guarantees uniqueness of the solution to the SDE, whereas the growth condition ensures that it does not explode in finite time. For most of our applications (many of which involve discrete-time approximations to the forward rate processes)

uniqueness is not a particularly critical feature and we shall shortly look at a process that violates the local Lipschitz condition at zero. The technical problems that arise are typically inconsequential and can, as we shall see, be taken care of in various ad-hoc ways if necessary. The less restrictive Yamada and Watanabe condition (Karatzas and Shreve (1991), p.291), can be shown to provide no weakening of the conditions in Theorem 1 for SDEs for F_k in measures other than $\mathbf{Q}^{k+1}$.

Provided that j is regular enough to allow for a unique, non-negative solution to the no-arbitrage forward rate SDEs (for instance, by satisfying the conditions of Theorem 1), we will refer to (4) as the *extended market model*. More general than the log-normal approach, the extended market model still remains quite tractable, particularly when it comes to the pricing of caps and floors. The following Lemma is useful for exploiting the "separable" form of (4) through a time-change:

Lemma 1.
Define

$$v_k(t) = \int_0^t \| l_k(u)\|^2 \, du \,, \qquad \tilde{W}_{k+1}(v_k(t)) = \int_0^t \| l_k(u)\| dW_{k+1}(u) \,.$$

$\tilde{W}_{k+1}$ *is a m-dimensional Brownian motion under the deterministic time-change* $v_k(t)$ *and* (1) *can be represented as the SDE*

$$df_k(v_k(t)) = j\left(f_k(v_k(t))\right) d\tilde{Z}^k_{k+1}(v_k(t)) \,, \tag{5}$$

where $f_k(v_k(t)) = F_k(t)$, *and* $\tilde{Z}^k_{k+1}(v_k(t)) \equiv l_k(t)^{\mathrm{T}} / \| l_k(t)\| \tilde{W}_{k+1}(v_k(t))$ *is a one-dimensional Brownian motion.*

Proof:
Follows from standard results for time-changes of Brownian motions, see e.g. Øksendahl (1995, p. 141).♠

A natural way to calibrate the extended market model is to parametrize j directly and back out ("imply") the l_k-functions from generic options (see e.g. Brace *et al* (1997) and Sidenius (1997) for various approaches). As we shall see, the fact that, in some sense, the l_k-functions can be collapsed into a time-change will greatly facilitate the process of constructing these functions from market data.

3.1. Case study: the CEV process

An example of a specification of a separable model is

$$j(x) = x^{a} \tag{6}$$

where a is a positive constant. This is the constant elasticity of variance (CEV) model studied in the context of equity option pricing by Cox and Ross (1976) and Schroder (1989).

The CEV model specification will *not* satisfy the local Lipshitz condition of Theorem 1 for $0 < a < 1$, and will violate the linear growth condition for $a > 1$. Theorem 1 can therefore not be used to characterize the properties of the market model with the CEV specification of $j(\cdot)$. As the CEV process is attractive both empirically and theoretically, we have compiled some useful results about the CEV model below:

Lemma 2.
Consider the stochastic differential equation

$$dx(v) = x(v)^{a}\, dZ(v), \tag{7}$$

where a is a positive constant and Z is a one-dimensional Brownian motion. The following holds:

a) *All solutions to* (7) *are non-explosive.*
b) *For $a \geq 1/2$ the SDE* (7) *has a unique solution.*
c) *For $0 < a < 1$, $x = 0$ is an attainable boundary for the process* (7)*; for $a \geq 1$, $x = 0$ is an unattainable boundary for the process* (7).
d) *For $0 < a < 1/2$ the SDE* (7) *does not have a unique solution, unless a separate boundary condition is specified for the boundary behavior in $x = 0$.*

Proof:
In Appendix A.♠

For $\frac{1}{2} \leq a < 1$, results b) and c) in Lemma 2 combined with the time-change representation in Lemma 1 implies that the SDE

$$dF_k(t) = F_k(t)^{a}\; l_k^{\mathrm{T}}(t) dW_{k+1}(t) \tag{8}$$

gives rise to a naturally occurring absorbing barrier in $F_k = 0$. According to d), however, if $0 < a < 1/2$ the behavior of F_k in 0 is not unique and requires us to select between the possible solutions. In our case, the choice of boundary condition is dictated by the no-arbitrage condition which requires that F_k remains a martingale (not just a *local* martingale), even when started at 0. It is thus clear that when $0 < a < 1/2$ we must insist on 0 being an *absorbing boundary* for F_k. In total, we thus associate (8) with an absorbing boundary at $F_k = 0$ for all $0 < a < 1$.

With 0 always being an absorbing barrier, the transition density of the process (8) can be written down in closed form. The result is stated in Lemma 3 for later use in the discussion of caplet pricing.

Lemma 3.
Consider the SDE (8) *for positive* $a \neq 1$, *and define*

$$X_k(T) = \frac{F_k(T)^{2(1-a)}}{(1-a)^2},\ J = -\frac{1}{2(1-a)},\ v_k(t,T) = \int_t^T \| l_k(u) \|^2 du,$$
$$I_a(x) = \sum_{j=0}^{\infty} \frac{(x/2)^{a+2j}}{j!\Gamma(a+j+1)},\ \Gamma(x) = \int_0^\infty u^{x-1} e^{-u} du\,.$$

Let $q_{k+1}(X_k(T)|X_k(t))$ *be the conditional density of* $X_k(T)$ *given* $X_k(t)$, $t \le T$, *under the probability measure* $\mathbf{Q}_{k+1}$. *If the level* $F_k = 0$ *is defined to be an absorbing boundary for* (8) *when* $0 < a < 1/2$, *then, for all positive* $a \neq 1$,

$$q_{k+1}(X_k(T)|X_k(t)) = \frac{1}{2v_k(T)} e^{-\frac{X_k(T)-X_k(t)}{2v_k(t,T)}} \left(\frac{X_k(T)}{X_k(t)}\right)^{-J/2} I_{|J|}\left(\frac{\sqrt{X_k(t)X_k(T)}}{v_k(t,T)}\right).$$

Proof of Lemma 3:
In Appendix A.♠

4. Caplet pricing

To calibrate the extended market model to market, it is important that efficient algorithms for the pricing of generic, liquid instruments be available. In this section we will consider the pricing of interest rate *caplets*, that is, instruments that at time T_{k+1} pay the amount $d_k(F_k(T_k) - H)^+$. While closed-form expressions are available in certain instances (see the later case study on the CEV process), we generally will need to rely on numerical methods. As we shall see, applying such methods as part of a calibration process is not as computationally demanding as it might seem.

Theorem 2.

Let $C_k(t)$ *denote the price of a Libor caplet maturing at time* T_k *with strike* H *and payment time* T_{k+1}. *Assume that forward rate dynamics satisfy (4). In the absence of arbitrage,* $C_k(t)$ *is given by*

$$C_k(t) = d_k P(t,T_{k+1}) g\big(v_k(t,T_k), F_k(t)\big), \quad v_k(t,T_k) = \int_t^{T_k} \| l_k(u) \|^2 du,$$

where $g(t,x)$ *solves the initial-value problem*

$$-\frac{\partial g}{\partial t} + \tfrac{1}{2} j(x)^2 \frac{\partial^2 g}{\partial x^2} = 0, \qquad g(0,x) = (x-H)^+. \tag{9}$$

Proof:

In Appendix A.♠

To solve the PDE (9) numerically, we can, for example, use a Crank-Nicholson finite difference scheme (e.g. Smith (1985)). Appendix B briefly discusses the mechanics of this scheme and verifies that a direct discretization of (9) is stable and convergent. Occasionally, one can take advantage of special forms of j and introduce transformations of x to improve the properties of the finite difference scheme. For example, when $j(x) = x$ it is customary (and appropriate) to introduce $y = \ln x$ and discretize in y. For everywhere differentiable j, the transformation

$$y(x) = \int j(x)^{-1} dx \tag{10}$$

might offer numerical advantages over a direct discretization (see e.g. Jamshidian (1991)) provided, of course, that the inverse of (10) exists and can be computed in closed form.

A crucial point about (9) is that the functions $g(.,.)$ only depend on the strike H and is independent of the initial forward rate $F_k(t)$ as well as the function $l_k(t)$. This means that we can use the same finite difference grid to price caplets with different maturities (and thus different forwards and volatilities), as long as the strikes remain the same. For example, the price of the T_l-maturity caplet with strike H is given by $C_l(t) = d_l P(t,T_{l+1}) g\big(v_l(t,T_l), F_l(t)\big)$ which, in finite difference terms, just corresponds to another grid-cell than the one used to pick up $C_k(t)$. In general, to solve for caplet prices maturing at all T_j, $n(t) \le j \le K$, we would need as many finite difference lattices as there are different strikes[3]. More importantly, in a model calibration where

[3] An exception occurs when $j(x)$ is a power function. Here, only a single grid is needed as we can normalize all strikes to a common number and absorb the normalization constants into the definition of v_k. For power functions,

we are running some root-search algorithm to determine the functions $l_k(t)$, we only need to compute these finite difference grids *once*, before the search routine is launched. This is so, because for each relevant caplet the finite difference grid will generate a vector of caplet prices in the t-domain. For each iteration on $l_k(t)$, determining the corresponding caplet prices thus becomes a mere matter of computing $v_k(t,T_k)$ and looking up (possibly interpolating) the right t-entry in the finite difference grid.

As a final point, notice that if the number of strikes exceeds the number of initial forwards, it is computationally advantageous to replace the backward equation (9) with the forward equation of Dupire (1994). In this approach, calendar time *t* and the initial forward are considered fixed, whereas caplet maturity and strike are variable, i.e. $C_k = C(T_k, H; t, F_k(t))$. Writing $C_k(t) = d_k P(t,T_k) h\big(0,H;v_k(t,T_k),F_k(t)\big)$, $h(t,x)$ solves the forward PDE

$$-\frac{\partial h}{\partial t} + \tfrac{1}{2}\sigma(x)^2 \frac{\partial^2 h}{\partial x^2} = 0 \qquad (11)$$

subject to the boundary condition $h(0,x;0,F_k(t)) = (F_k(t)-x)^+$. (11) can be discretized using the same approach as for (9). We need one finite difference grid for all different time *t* forwards, but each grid can accommodate different strikes and maturities.

4.1. Case study: the CEV process

For the CEV specification (6) studied in the previous section it is possible to obtain closed-form expressions for caplet prices. The results are contained in Theorem 3 below.

Theorem 3.

As above, let $C_k(t)$ denote the arbitrage-free price of a Libor caplet maturing at time T_k with strike H and payment time T_{k+1}. Also, let $N(\cdot)$ be the standard Normal cumulative distribution function, and $\chi^2(\cdot, J, \lambda)$ be the cumulative distribution function for a non-central χ^2-distributed random variable with non-centrality parameter λ and J degrees of freedom. Define

$$a = \frac{H^{2(1-\alpha)}}{(1-\alpha)^2 v_k(t,T_k)},\ b = \frac{1}{1-\alpha},\ c = \frac{F_k(t)^{2(1-\alpha)}}{(1-\alpha)^2 v_k(t,T_k)},$$

$$x_\pm = \frac{\ln[F_k(t)/H] \pm \tfrac{1}{2} v_k(t,T_k)}{\sqrt{v_k(t,T_k)}},\quad v_k(t,T_k) = \int_t^{T_k} \| \sigma_k(u) \|^2 du\ .$$

---

however, the closed-form caplet pricing solution in Theorem 3 is more convenient than the finite difference method.

Assuming that forward rates evolve according to (8), we then have the following results:

a) *For* $0 < a < 1$ *and an absorbing boundary at the level* $F_k = 0$*:*

$$C_k(t) = d_k P(t, T_{k+1})\left[F_k(t)\left(1 - c^2(a, b+2, c)\right) - Hc^2(c, b, a)\right].$$

b) *For* $a = 1$:

$$C_k(t) = d_k P(t, T_{k+1})\left[F_k(t) N(x_+) - HN(x_-)\right].$$

c) *For* $a > 1$*:*

$$C_k(t) = d_k P(t, T_{k+1})\left[F_k(t)\left(1 - c^2(c, -b, a)\right) - Hc^2(a, 2-b, c)\right].$$

Proof:
b) is just the usual Black (1976) caplet formula for log-normal forward rates. a) and c) are proven in Appendix A.♠

The non-central c^2-distribution function in the above caplet formulas can be computed using, for instance, the efficient numerical procedure described by Ding (1992).

Below we report implied (Black (1976) formula) log-normal volatilities for the cases $a = 0.5,\ l_k = 0.05$ and $a = 1.5,\ l_k = 0.83$, for different strikes and maturities. In both cases, the l_k's are assumed constant as a function of time and *k*. Also, the initial forward curve is assumed flat at $F_k = 0.06$ for all k.

Implied log-normal volatilities for CEV caplet prices
$a = 0.5,\ l_k = 0.05,\ F_k = 0.06$

$T_k \backslash H$	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
1	0.3109	0.2655	0.2417	0.2256	0.2137	0.2042	0.1964	0.1899	0.1842	0.1792
5	0.3104	0.2662	0.2422	0.2261	0.2141	0.2046	0.1967	0.1901	0.1844	0.1795
10	0.3114	0.2670	0.2428	0.2266	0.2145	0.2049	0.1971	0.1905	0.1847	0.1797
20	0.3113	0.2677	0.2436	0.2273	0.2152	0.2056	0.1977	0.1910	0.1853	0.1802
30	0.3084	0.2670	0.2436	0.2276	0.2155	0.2059	0.1981	0.1914	0.1856	0.1806

Table 1A

Implied log-normal volatilities for CEV caplet prices

$a = 1.5,\ l_k = 0.83,\ F_k = 0.06$

$T_k \backslash H$	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
1	0.1527	0.1702	0.1835	0.1943	0.2034	0.2113	0.2184	0.2247	0.2305
5	0.1527	0.1704	0.1837	0.1946	0.2037	0.2117	0.2188	0.2252	0.2310
10	0.1529	0.1706	0.1840	0.1949	0.2041	0.2121	0.2193	0.2257	0.2315
20	0.1532	0.1710	0.1845	0.1955	0.2047	0.2128	0.2200	0.2264	0.2323
30	0.1535	0.1714	0.1849	0.1958	0.2051	0.2132	0.2203	0.2267	0.2324

Table 1B

As is evident from the tables, by varying the parameter a , the CEV process can generate both downward-sloping and upward-sloping volatility skews. The figure below emphasizes this point by graphing the implied 3-year caplet volatility skew for various values of a . In the figure, the forward curve is constant at $F_k = 0.06$ and l_k (independent of time) is for each a set such that the at-the-money (H = 0.06) implied log-normal volatility equals 30%.

Implied log-normal caplet volatilities as a function of CEV a

$T_k - t = 3;\ F_k = 0.06$

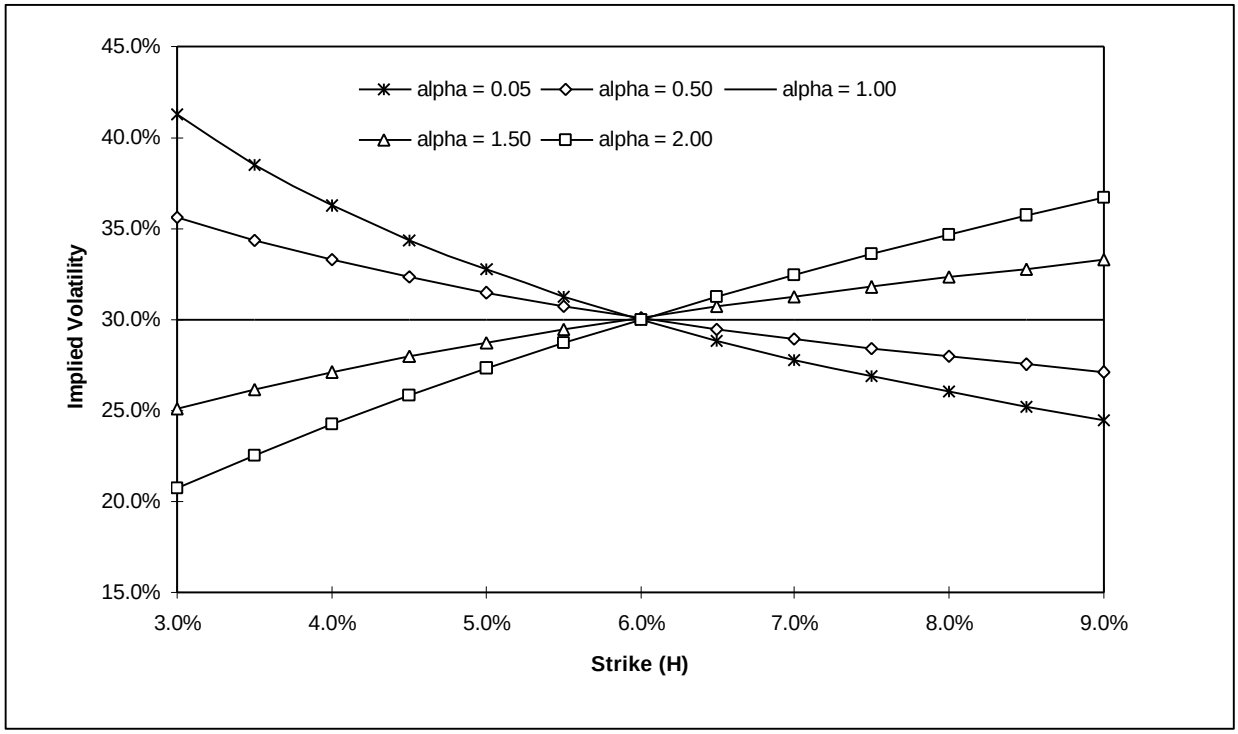

Figure 2

4.2. Case Study: the 'Limited' CEV (LCEV) Process

One problem with the CEV specification of the market model is that for $a < 1$, the origin is an attainable and absorbing boundary for the process. The positive probability of absorbed forwards is not necessarily a problem for the pricing of caplets, but is obviously not desirable from an empirical standpoint and might also create some difficulties in the pricing of more exotic structures. On the other hand, when $a > 1$ the growth condition of Theorem 1 is not satisfied and that could well create problems with exploding interest rate paths under probability measures where the forward drift is non-zero.

To overcome the regularity problems of the CEV process, we can specify a 'regularized' version of CEV market model by letting

$$j(x) = x \cdot \min\left(e^{a-1}, x^{a-1}\right), \quad e > 0 \tag{12}$$

where e is a small fixed number when $a < 1$, and a large fixed number when $a > 1$. We will use the term *limited CEV* (*LCEV*) *process* for the specification (12). Roughly speaking, when forward rates cross over the 'switching level' e, the LCEV process becomes a geometric Brownian motion with a high, but finite, volatility. Notice that the LCEV process satisfies the Lipschitz and growth conditions of Theorem 1 and that zero is an unattainable boundary for all F_k. The drawback of the (12) is, of course, that the closed-form expressions for caplets can only be used as approximations and that exact caplet prices must, in theory at least, be obtained by numerical procedures. As we shall see, however, the cap price dependence on e is typically limited, making the closed-form expressions in Theorem 3 sufficient for calibration purposes.

To obtain caplet prices under the LCEV process, we first turn to the finite difference scheme outlined earlier. In the tables below, we give implied log-normal volatilities for the two scenarios we considered earlier for the CEV process. All numbers were generated using a Crank-Nicholson finite difference scheme with a mesh size of 500×500 points.

Implied log-normal volatilities for LCEV caplet prices (500 x 500 CN scheme)

$a = 0.5,\ l_k = 0.05, e = 0.0025,\ F_k = 0.06$

T\H	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
1	0.3099	0.2657	0.2417	0.2256	0.2137	0.2042	0.1964	0.1899	0.1842	0.1792
5	0.3104	0.2662	0.2422	0.2261	0.2141	0.2045	0.1967	0.1901	0.1844	0.1795
10	0.3113	0.2670	0.2428	0.2266	0.2145	0.2049	0.1971	0.1905	0.1847	0.1797
20	0.3109	0.2676	0.2436	0.2273	0.2152	0.2056	0.1977	0.1910	0.1853	0.1802
30	0.3078	0.2668	0.2435	0.2275	0.2155	0.2059	0.1981	0.1914	0.1856	0.1806

Table 2A

Implied log-normal volatilities for LCEV caplet prices (500 x 500 CN scheme)

$a = 1.5,\ l_k = 0.83, e = 0.25,\ F_k = 0.06$

T\H	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
1	0.1532	0.1706	0.1835	0.1943	0.2034	0.2113	0.2184	0.2247	0.2305
5	0.1532	0.1705	0.1837	0.1945	0.2037	0.2117	0.2188	0.2252	0.2310
10	0.1532	0.1706	0.1840	0.1949	0.2041	0.2121	0.2192	0.2256	0.2314
20	0.1534	0.1710	0.1845	0.1954	0.2045	0.2124	0.2194	0.2255	0.2309
30	0.1534	0.1713	0.1846	0.1954	0.2043	0.2119	0.2184	0.2242	0.2292

Table 2B

Comparing Table 2A-B to the results for the unconstrained CEV case (Table 1A-B), we see that the closed-form formula is a good approximation for the LCEV cap prices, although the choices of e above result in a slight rounding of the corners of the implied volatility grids.

By letting either $e \to 0_+$ ($a < 1$) or $e \to \infty$ ($a > 1$), one would expect the LCEV process to approach the CEV process. This is formalized in the Theorem below:

Theorem 4.

Suppose that

$$dx(v) = x^a(v)dZ(v),$$
$$dy(v) = \min\left(e^{a-1}, y(v)^{a-1}\right)y(v)dZ(v),$$

where $a > 0$, *and* $x(0) = y(0) > 0$. *For* $0 < a < \frac{1}{2}$, *0 is an absorbing boundary for x. Let* $h > 0$, $H > 0$, *and* $T > 0$ *be given, and let* **P** *be the relevant probability measure. Then*

a) *For* $a \geq 1$*:*

$$\lim_{e\to\infty}\left|\mathbf{P}\left(x(T) < h\right) - \mathbf{P}\left(y(T) < h\right)\right| = 0,$$
$$\lim_{e\to\infty}\left|E^{\mathbf{P}}\left[(x(T) - H)^+\right] - E^{\mathbf{P}}\left[(y(T) - H)^+\right]\right| = 0.$$

b) *For* $a < 1$*:*

$$\lim_{e\to 0_+}\left|\mathbf{P}\left(x(T) < h\right) - \mathbf{P}\left(y(T) < h\right)\right| = 0,$$
$$\lim_{e\to 0+}\left|E^{\mathbf{P}}\left[(x(T) - H)^+\right] - E^{\mathbf{P}}\left[(y(T) - H)^+\right]\right| = 0.$$

Proof:

In Appendix A.♠

Theorem 3 shows that prices of caps and (by put-call parity) floors in the CEV and LCEV models can be brought to converge to each other. To investigate the speed of this convergence, consider the computation of $C_k(t)/P(t,T_{k+1}) = E_t^{\kappa+1}\left[(F_k(T_k)-H)^+\right]$ when $T_k - t = 30$, $F_k(t) = 0.02$, $a = 0.5$, and $l_k = 0.09$ (independent of time). The low rates, long maturity, and high volatility (around 65% in log-normal terms) implies a probability of eventual absorption of 84.8% (!) in the CEV process, so our example should emphasize the differences between the CEV and LCEV processes. In Table 3, we have used the Crank-Nicholson scheme to compute $C_k(t)/P(t,T_{k+1})$ in the LCEV process, for various values of e and *H*. The table also contains the exact CEV value based on formula a) in Theorem 2.

30-year undiscounted caplet prices (in basis points) for LCEV model (500 x 500 CN)

$T_k - t = 30$, $F_k(t) = 0.02$, $a = 0.5$, and $l_k = 0.09$

e / *H*	0.01	0.0125	0.015	0.02	0.025	0.03	0.04	0.05
0.005	185.1828	181.6836	178.2509	171.5790	165.1633	158.9841	147.3179	136.5103
0.004	185.2726	181.7794	178.3522	171.6905	165.2824	159.1094	147.4516	136.6487
0.003	185.3204	181.8327	178.4096	171.7525	165.3489	159.1795	147.5269	136.7268
0.002	185.3637	181.8782	178.4575	171.8054	165.4056	159.2394	147.5911	136.7937
0.001	185.3849	181.9003	178.4807	171.8313	165.4335	159.2688	147.6228	136.8266
0.0005	185.3841	181.9004	178.4811	171.8314	165.4335	159.2688	147.6228	136.8267
0.00025	185.3835	181.9003	178.4813	171.8314	165.4335	159.2688	147.6229	136.8267
0.0001	185.3830	181.9001	178.4813	171.8314	165.4336	159.2689	147.6229	136.8267
0	185.3827	181.9000	178.4814	171.8314	165.4336	159.2689	147.6229	136.8268
CEV	185.3829	181.8984	178.4791	171.8319	165.4317	159.2694	147.6235	136.8273

Table 3

Despite the extreme parameters of the option in Table 3, the prices generated by the LCEV model are very close to those of the CEV model. In fact, for no combination of e and *H* in the table above is the price difference larger than 0.4 bp. Using the CEV caplet pricing formula to price caplets under a reasonably truncated LCEV process is generally justifiable and will help improve the speed at which the LCEV process can be calibrated to cap market quotes.

5. Swaption pricing

In Section 4, we saw that the properties of the extended market models makes the pricing of caps and floors straightforward, a feature that greatly facilitates the calibration of the model to quoted market prices of these instruments. While cap and floor prices might sometimes suffice for model calibration (see e.g. Jamshidian (1997) for a discussion), one would, however, normally also want to supplement the calibration with quoted prices of at-the-money (ATM) swaptions.

Unfortunately, the Libor market model framework does not allow for an exact closed-form swaption pricing formula. Jamshidian (1997) discusses the usage of Monte Carlo simulation to price swaptions in the Libor market model, although he realizes the computational problems in embedding simulations in the calibration root-search algorithm[4]. We discuss Monte Carlo pricing of swaptions and other fixed income derivatives in section 6.

A more attractive approach is suggested in Brace *et al* (1997) who develop a closed-form (log-normal) approximation that appears to work reasonably well for a range of market and swaption parameters. Briefly speaking, Brace *et al* make a rank-1 assumption about a certain variance-covariance matrix and approximate a series of (stochastic) forward rate drifts as deterministic functions of time. These assumptions allow for a decomposition along the lines of Jamshidian (1989) and necessitates root-search algorithms to locate the critical level of a Gaussian perturbation factor. While the approximations in Brace *et al* can be extended to cover Gaussian models, they are not easily expanded to more complicated processes. In this section, we seek to develop analytical swaption price approximations that are broad enough for the models considered in this paper. Our approximations are, essentially, based on par rate dynamics in a forward swap measure, and result in pricng PDEs that are of the same type as for caplets (Section 4). While it is difficult to construct accurate approximations for every imaginable form and level of forward rate volatilities, our approach seems to work well for most "reasonable" specifications. For the log-normal and CEV specifications of the extended market model, closed-form solutions are available that are are both faster (no root-search necessary) and, in our experience, often more accurate than the formulas in Brace *et al*. Section 6 contains some tests of our approximations using the CEV / LCEV model as an example.

Consider now a European payer swaption maturing at some date T_s, $s \in \{1,2,\ldots,K\}$. The swaption gives the holder the right to pay fixed cashflows[5] $q\,d_{k-1} > 0$ at T_k, for $k = s+1, s+2, \ldots, e$ in exchange for 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 the payer swaption *S* is, by definition,

$$S(T_s) = \left(\sum_{k=s}^{e-1} P(T_s, T_{k+1})\, d_k [F_k(T_s) - q] \right)^+ = \left(1 - P(T_s, T_e) - q \sum_{k=s}^{e-1} d_k P(T_s, T_{k+1}) \right)^+ . \qquad (13)$$

[4] In defense of the Monte Carlo method, we should point out, however, that since each Monte Carlo path allows for simultaneous pricing of all swaptions in the calibration set, calibration will typically still be computationally feasible, albeit slow.

[5] In practice, the fixed side of the swap might sometimes have a different payment basis and/or frequency than the floating side. Handling of arbitrary fixed cash-flow streams can be done by changing the definition of the numeraire in (14) to include the actual cash-flows. For notational convenience, we have omitted this simple extension here.

We now define

$$B^S(t) = \sum_{k=s}^{e-1} d_k P(t, T_{k+1}), \tag{14}$$

$$R(t) = \frac{P(t, T_s) - P(t, T_e)}{B^S(t)}, \tag{15}$$

whereby we can write (13) in the alternative form

$$S(T_s) = B^S(T_s)\big(R(T_s) - q\big)^+. \tag{16}$$

$B^S(t)$ in (14) is a strictly positive process and can thus be used as a pricing numeraire. We use $\mathbf{Q}^S$ to denote the measure induced by this numeraire; we will refer to $\mathbf{Q}^S$ as the *forward swap measure* for the swaption *S*. By standard theory, absence of arbitrage implies that the so-called *par-rate* $R(t)$ in (15) is a martingale under $\mathbf{Q}^S$. Assuming that the yield curve dynamics are governed by an extended market model of the type (4), an application of Ito's lemma to (15) yields,

$$dR(t) = \sum_{j=s}^{e-1} \frac{\partial R(t)}{\partial F_j} \jmath\left(F_j(t)\right) \lambda_j(t)^{\mathrm{T}} dW^S(t),$$

where W^S is a *m*-dimensional Brownian motion under $\mathbf{Q}^S$ and

$$\frac{\partial R(t)}{\partial F_j} = \frac{R(t)\, d_j}{1 + d_j F_j(t)} \left[\frac{P(t, T_e)}{P(t, T_s) - P(t, T_e)} + \frac{\sum_{k=j}^{e-1} d_k P(t, T_{k+1})}{B^S(t)} \right]. \tag{17}$$

Due to the complexity of (17), the process for *R*(*t*) is not tractable. To proceed, we notice that for most reasonable shifts of the forward curve, the expression (17) normally varies very little with time and the state of interest rates. For forward curve movements that are predominantly parallel (which is the case in practice), it is often also reasonable to assume that the ratio

$$\jmath\left(F_j(t)\right) / \jmath\left(R(t)\right)$$

is close to constant. In total, we suggest the following approximation of the dynamics of R in the extended Libor market model:

$$dR(u) \approx j\left(R(u)\right)\sum_{j=s}^{e-1} w_j(t)\, l_j^{\mathrm{T}}(u) dW^S(u),\ \ t \le u < T_e\,, \tag{18}$$

where

$$w_j(t) = \frac{\partial R(t)}{\partial F_j(t)} \frac{j\left(F_j(t)\right)}{j\left(R(t)\right)} \tag{19}$$

can be computed from (17).

With the approximation (18), the SDE for R under $\mathbf{Q}^S$ takes exactly the same form as the forward rate SDE's discussed in Section 3, and all the Theorems and Lemmas of this section apply. Further, the form of (16) shows that the swaption pricing problem becomes identical (after substitution of numeraires) to the caplet pricing problem discussed in detail in Section 4. From (9), for instance, we can write the swaption price as

$$S(t) = B^S(t)\, p\big(v_S(t,T_s), R(t)\big),\ \ v_S(t,T_s) = \int_t^{T_s} \left\| \sum_{j=s}^{e-1} w_j(t)\, l_j(u) \right\|^2 du,$$

where $p(t,x)$ solves

$$-\frac{\partial p}{\partial t} + \tfrac{1}{2} j\,(x)^2 \frac{\partial^2 p}{\partial x^2} = 0\,, \qquad g(0,x) = (x-q)^+\,.$$

In a calibration, we would need to solve as many PDEs as there are different swaption coupons; as discussed in Section 4, however, these PDEs only must be solved once, before the calibration loop is started.

As for caplets, the CEV specification of j allows for a closed-form pricing formulas, listed below for convenience. The formulas also serve as an approximation for swaption prices under the LCEV forward rate process.

Theorem 5.

Consider a European payer swaption on a swap with start date T_s, *end date* T_e, *and fixed cashflows* $q\,d_{k-1} > 0$ *at* T_k, *for* $k = s+1,\ldots,e$. *Assume that the forward rate dynamics are given by the CEV specification (8) and that the approximation (18) holds. Define*

$$d = \frac{q^{2(1-a)}}{(1-a)^2 v_S(t,T_s)},\ b = \frac{1}{1-a},\ f = \frac{R(t)^{2(1-a)}}{(1-a)^2 v_S(t,T_s)},$$

$$g_{\pm} = \frac{\ln[R(t)/q] \pm \frac{1}{2} v_s}{\sqrt{v_s}}, v_S(t,T_s) = \int_t^{T_s} \left\| \sum_{j=s}^{e-1} w_j(t)\, l_j(u) \right\|^2 du,$$

where $R(t)$ *is defined in (15) and the* w_j*'s in (17) and (19). Also let* $B^S(t)$ *be as in (14). In the absence of arbitrage, the price at time* $t \le T_s$ *of the swaption is given by*

a) *For* $0 < a < 1$ *and an absorbing boundary at the level* $F_k = 0$*:*

$$S(t) = B^S(t)\left[R(t)\left(1 - c^2(d, b+2, f)\right) - q c^2(f, b, d)\right].$$

b) *For* $a = 1$*:*

$$S(t) = B^S(t)\left[R(t) N(g_+) - q N(g_-)\right]$$

c) *For* $a > 1$*:*

$$S(t) = B^S(t)\left[R(t)\left(1 - c^2(f, -b, d)\right) - q c^2(d, 2-b, f)\right].$$

Proof:

Follows directly from the proof of Theorem 4.♠

6. Monte Carlo simulation

Having discussed efficient ways to price the simple instruments that can form the basis for a market calibration, we now turn to application of the extended market model to the pricing of more complicated OTC structures. Consider a final maturity $T \le T_{K+1}$ and a derivative security with payout $V(T)$ at time *T*. $V(T)$ is allowed to depend on the path of all forwards $F_k(t)$, $k = 0,1,\ldots,K$, for $t \in [0,T]$. Under the spot measure **Q**, we have

$$V(0) = E^{\mathbb{Q}}\left[V(T)/B(T)\right] \tag{20}$$

where $E^{\mathbb{Q}}[\cdot]$ denotes expectation under the spot measure and the numeraire B is defined in (3). Evaluation of (20) can very rarely be done analytically, hence we need to consider numerical methods. As the forward rate evolution in the extended market model is generally non-Markov and involves multiple factors, trees and lattices are typically not computationally feasible (see for example discussion in Jamshidian (1997)) and we here exclusively deal with the Monte Carlo method. We will assume that the reader is familiar with the basics of this technique; a good survey of Monte Carlo methods in finance is Boyle *et al* (1997). As indicated in (20), we will work in the spot measure throughout, but point out that other measures would do as well.

To generate random paths of forward rates, we first set up a L-dimensional simulation time grid, $0 = t_0 < t_1 < ... < t_{L-1} < t_L = T$. Being general, we do not require this grid to subdivide (or equate) the maturity grid $0 = T_0 < T_1 < ... < T_{K+1}$, but we point out that it is often convenient to at least have the maturity grid be a subset of the simulation grid. The separation of maturity and simulation grids makes some computations a bit more cumbersome, but allows us to maintain a steady, equidistant forward rate maturity structure while at the same time allowing for perfect alignment of the simulation time grid with all dates required in the payout computations[6]. Given the finite set of simulation dates, we obviously have no hope of simulating the continuous-time processes F_k, but must contend ourselves with some approximation $\hat{F}_k$ defined on $\{t_0, t_1, ..., t_L\}$.

Consider now a specific date t_i and assume that all $\hat{F}_k(t_i)$, $n(t_i) \le k \le K$, are known. The simplest way to advance the simulation to t_{i+1} is to apply an *Euler scheme* to the continuous-time SDE (2):

$$\hat{F}_k(t_{i+1}) = \hat{F}_k(t_i) + j\left(\hat{F}_k(t_i)\right) l_k^T(t_i)\left[\hat{m}_k(t_i)\Delta_i + e_i\sqrt{\Delta_i}\right], \quad n(t_{i+1}) \le k \le K \tag{21}$$

where $\Delta_i \equiv t_{i+1} - t_i$, e_i is a m-dimensional vector of independent standard Gaussian variables, and

$$\hat{m}_k(t_i) = \sum_{j=n(t_i)}^{k} l_j(t_i)\frac{d_j j\left(F_j(t_i)\right)}{1 + d_j F_j(t_i)}. \tag{22}$$

In the equation (22) for $\hat{m}_k(t_i)$, there is some ambiguity in the choice of the lower summation index if t_i falls on a date in the maturity structure, say $t_i = T_a$ for some integer a. Due to our

[6] The generality of our setup will quite frequently require the pricing of zero coupon bonds that do not mature directly on a date in the maturity structure. This can be accomplished through one of various possible interpolation techniques. Similarly, one will typically need schemes to interpolate (or extrapolate) $P(t, n(t))$ when t is not in the maturity structure.

definition of the mapping function $n(\cdot)$ as being right-continuous, we exclude the term $l_a(T_a)\,d_a j\left(F_a(T_a)\right)/(1+d_a F_a(T_a))$ in $\hat{m}_k(t_i)$. In continuous time, the term would disappear at $t_i + dt$, so on intuitive grounds it is reasonable to ignore it and, in effect, start the sum from index $a+1$. As was the case for finite difference schemes, it is sometimes advantageous to write down Euler schemes for some suitable chosen transformation of forward rates (see (10)). For instance, the Euler discretization of the evolution of log-transformed forward rates leads to the multiplicative simulation scheme below:

$$\hat{F}_k(t_{i+1}) = \hat{F}_k(t_i)\exp\left(\frac{j\left(\hat{F}_k(t_i)\right)}{\hat{F}_k(t_i)}\; l_k^T(t_i)\left[\left(\hat{m}_k(t_i) - \tfrac{1}{2}\frac{j\left(\hat{F}_k(t_i)\right)}{\hat{F}_k(t_i)}\; l_k(t_i)\right)\Delta_i + e_i\sqrt{\Delta_i}\right]\right) \qquad (23)$$

Unlike (21), (23) guarantees that forward rates are positive. As such, (23) should strictly speaking not be used for models where zero is attainable (such as the CEV process).

Independent of the form and magnitude of j and l, the discrete-time dynamics of the direct Euler scheme (21) will always involve a finite probability of generating negative forward rates, even if such rates are unattainable in a continuous-time setting. In some cases this can be ignored, but for processes where j is not defined for negative arguments (such as the CEV process introduced earlier) heuristic rules must be employed to ensure that the forwards stay non-negative. For instance, for CEV processes it is natural to absorb all negative forwards at zero. For processes where 0 is known to be unattainable, negative forward rates can be 'reflected' at zero and be replaced by their absolute values.

The Euler schemes (21) and (23) are simple and easy to implement, but have a built-in bias relative to the true continuous-time SDE (2). That is, only in the limit $\Delta_i \to 0$ would the simulated dynamics of forward rates match the continuous-time distribution implied by the SDE (2). While the Euler scheme will eventually converge to the true distribution, the speed of convergence is only of (weak) order 1 (see Kloeden and Platen (1992)). The presence of the first-order bias has two related implications. First, arbitrages amongst the various zero-coupon bonds in the maturity structure exist and the initial bond prices, $P(0,T_k)$, will not be replicated exactly, even for an infinite number of Monte Carlo trials. Second, the simulated prices of caps, floors and swaptions will exhibit a bias relative to the continuous-time prices (which we would normally use for model calibration, see Sections 4 and 5).

The first problem is normally the less severe of the two, and can sometimes be eliminated completely through suitable choice of simulation variable. For instance, if we can write down an unconstrained Euler scheme directly in $D_k(t) \equiv P(t,T_k)/B(t)$, the numeraire-deflated zero-coupon bond prices automatically become discrete-time martingales on $\{t_i\}$, resulting in bias-free prices of zero-coupon bonds. For processes with strictly non-negative state space (such as the

CEV process), unconstrained Euler discretization of $D_k(t)$ is, however, generally not possible but, as discussed earlier, must be supplemented with rules for dealing with occasional negative rates. These rules will destroy the discrete-time martingale property of the Euler scheme for $D_k(t)$. Glasserman and Zhao (1998) discuss discrete-time arbitrage-free schemes in much more detail, but as the biases in bond prices produced even by naïve schemes like (21) are typically very small, we shall not pursue the topic further here.

To address the problems of bias in more generality, one can turn to the methods of high-order discretization schemes; see for example the monograph by Kloeden and Platen (1992) for details. Unlike the technique discussed above, these schemes cannot eliminate biases completely, but merely increase the speed of convergence. On the other hand, the benefits of high-order discretization techniques are not limited to bond prices, but all (suitably regular) functionals of the path of forward rates, including option payouts. The classical example of a high-order simulation scheme is the 2nd-order Milstein-scheme (Kloeden and Platen (1992)) based on a stochastic Taylor expansion of forward rates. As it requires computation of explicit derivatives in time and forward rates, this scheme is, unfortunately, extremely cumbersome when applied to the $(K-n(t)+1)$-dimensional system of *m*-factor SDEs (2). (An exception occurs for the basic log-normal model with piecewise flat volatilities; the second-order Milstein scheme for this case can be found in Brotherton-Ratcliffe (1997)). A simpler way to apply a high-order simulation scheme is by *Romberg-Richardson extrapolation* (Talay and Tubaro (1990), Kloeden and Platen (1992)), where price estimates at different time-steps are combined to cancel off leading order error terms. Extrapolation schemes are simple and elegant, but they are typically less effective than schemes based on analytical derivatives. While we have not performed a systematic study, a number of Monte Carlo tests of extrapolation methods for caps and swaption prices in the CEV model gave disappointing results. For more details on extrapolation methods and some tests in a short-rate setting, see also Andersen (1995).

Before we proceed to test the suggested simulation algorithms on the CEV process, we should make the (obvious) point that the discrete-time Monte Carlo schemes dicussed above not only involve systematic biases, but are also subject to the usual random sample errors. Methods to control the variance of the sample error are surveyed in Boyle *et al* (1997) and, in an interest rate setting, Andersen (1995) and will not be discussed here.

6.1. Case Study: the square-root LCEV Process

In this section we will test and illustrate the method of Monte Carlo simulation by using it to compute the prices of bonds, caplets, and swaptions under the square-root LCEV dynamics (see Section 4). As we still wish to use the closed-form CEV caplet formula and swaption approximation, we set the parameter e in (12) to a low number, $e = 1/400$; that is,

$j(x) = x \cdot MIN\{20, x^{-1/2}\}$. $j(x) = x \cdot MIN\{20, x^{-1/2}\}$. We will assume that the forward curve has constant semi-annual spacing, i.e. $d_k = 0.5$ for all k, and consider the following two scenarios:

Scenario A:

Constant initial forward rates of 6%: $F_k(0) = 0.06$, for all k

Constant volatility of 5%: $l_k(t) = 0.05$, for all k and t

Scenario B:

Constant initial forward rates of 2%: $F_k(0) = 0.02$, for all k

Downward-sloping volatility function: $l_k(t) = MAX\{0.09 - 0.02(k - n(t)), 0.01\}$

The scenarios A and B are rough proxies for the current (summer 1998) market conditions in the US and Japan, respectively. Notice the extremeness of scenario B: in approximate log-normal terms, the initial volatilities start out at around 65% and fall to 25% for the 10-year forward. The time 0 caplet volatilities consistent with scenario B are displayed in Figure 3:

Caplet (square-root) Volatility vs. Caplet Maturity (Scenario B)

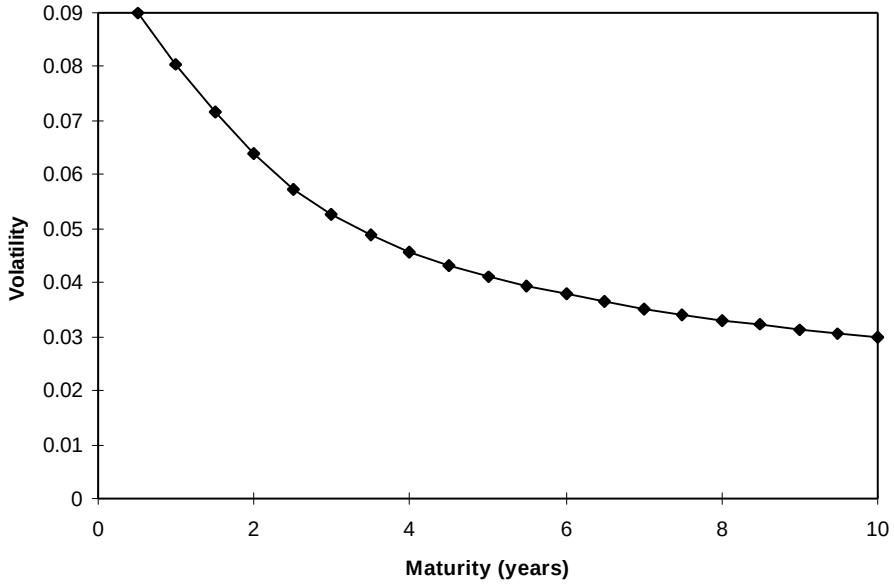

Figure 3

To simulate the LCEV process, we here use the log-Euler scheme (23) which ensures that the positivity of forward rates in the continuous-time LCEV process (see Theorem 1) is maintained in our simulations. For the square-root LCEV process, it also appears that the simulation bias of the

log-Euler scheme (23) is typically slightly lower than that of the Euler scheme (21). This might not be the case for arbitrary LCEV specifications, particularly not when a in (12) is low.

In Table 4A and 4B, we have used 1 million random paths (but no variance reduction techniques) in the Euler scheme (21) to estimate the prices of selected zero-coupon bonds under the scenarios above. The table contains sample price errors (simulated price - true price) and standard deviations (S.D.) for simulation time-steps of 0.5, 0.25, and 0.125.

Zero-coupon bond price simulation errors (in basis points) in scenario A
1,000,000 Monte Carlo Paths, Log-Euler Scheme, $j(x) = x \cdot MIN(20, x^{-1/2})$

Bond	$\Delta = 0.5$		$\Delta = 0.25$		$\Delta = 0.125$	
Maturity	Error	S.D.	Error	S.D.	Error	S.D.
1	0.000	0.001	0.000	0.001	0.001	0.002
2	0.01	0.01	0.00	0.01	-0.02	0.01
5	-0.03	0.03	-0.06	0.03	-0.06	0.04
7	-0.05	0.04	-0.04	0.04	-0.07	0.04
10	-0.05	0.07	-0.02	0.05	-0.06	0.05

Exact Prices: 0.9426 (1 yr); 0.8885 (2 yrs); 0.7441 (5 yrs); 0.6611 (7 yrs); 0.5537 (10 yrs)

Table 4A

Zero-Coupon Bond Price Simulation Errors (in basis points) in Scenario B
1,000,000 Monte Carlo Paths, Log-Euler Scheme, $j(x) = x \cdot MIN(20, x^{-1/2})$

Bond	$\Delta = 0.5$		$\Delta = 0.25$		$\Delta = 0.125$	
Maturity	Error	S.D.	Error	S.D.	Error	S.D.
1	0.03	0.01	0.03	0.01	-0.01	0.01
2	0.16	0.04	0.09	0.03	-0.05	0.03
5	0.30	0.08	0.16	0.06	-0.03	0.06
7	0.47	0.09	0.25	0.08	0.01	0.07
10	0.70	0.11	0.37	0.09	0.17	0.08

Exact Prices: 0.9803 (1 yr); 0.9610 (2 yrs); 0.9053 (5 yrs); 0.8700 (7 yrs); 0.8195 (10 yrs)

Table 4B

In the case of scenario A, all computed price errors are less than 0.1 bp. Moreover, as the price errors consistently are less than 2 sample standard deviations in magnitude, they are essentially indistinguishable from 0. In other words, while we know from the discussion earlier that the log-Euler scheme will result in a bond price bias, for the moderate volatilities in scenario A this bias is so small that not even 1,000,000 simulations allows us to separate it from random noise. In the

case of scenario B, however, the bias is still small, but can be distingusihed from noise, particularly for long maturities. For the 10-year bond, for instance, when $\Delta = 0.5$ the price bias amounts to 0.7 basis points or more than 6 standard deviations. As expected, the price errors in Table 4B fall when the simulation time-step is decreased. For the relatively coarse time-steps used, however, the convergence order generally appears higher than the expected number of 1.

Turning now to the pricing of caplets, tables 5A and 5B list our simulation results for at-the-money caplets of various maturities. The simulation errors in the tables are here defined as the simulated price minus the exact CEV price, the latter computed from Theorem 3, Section 4.

ATM Caplet Price Simulation Errors (in basis points) in Scenario A
1,000,000 Monte Carlo Paths, Log-Euler Scheme, $j(x) = x \cdot MIN(20, x^{-1/2})$

Caplet	$\Delta = 0.5$		$\Delta = 0.25$		$\Delta = 0.125$	
Maturity	Error	S.D.	Error	S.D.	Error	S.D.
1	0.002	0.02	0.006	0.02	0.05	0.02
2	0.03	0.03	0.03	0.03	0.08	0.03
5	0.05	0.03	0.04	0.03	0.08	0.03
7	0.001	0.03	-0.004	0.03	0.05	0.03
10	-0.05	0.03	-0.01	0.03	-0.02	0.03

Exact CEV Prices (bp): 22.33 (1 yr); 29.72 (2 yrs); 39.20 (5 yrs); 41.10 (7 yrs); 40.98 (10 yrs)

Table 5A

ATM Caplet Price Simulation Errors (in basis points) in Scenario B
1,000,000 Monte Carlo Paths, Log-Euler Scheme, $j(x) = x \cdot MIN(20, x^{-1/2})$

Caplet	$\Delta = 0.5$		$\Delta = 0.25$		$\Delta = 0.125$	
Maturity	Error	S.D.	Error	S.D.	Error	S.D.
1	0.12	0.02	0.15	0.02	0.16	0.02
2	0.23	0.03	0.20	0.03	0.17	0.02
5	0.22	0.03	0.20	0.03	0.15	0.02
7	0.20	0.02	0.17	0.03	0.08	0.02
10	0.21	0.02	0.20	0.03	0.04	0.02

Exact CEV Prices (bp): 21.85 (1 yr); 23.99 (2 yrs); 23.00 (5 yrs); 22.35 (7 yrs); 21.40 (10 yrs)

Table 5B

As was the case for zero-coupon bonds, simulated price errors for scenario A are mostly not statistically significant, not even for long caplet maturities. Somewhat surprisingly, the errors appear to be increasing with the number of simulation time-steps, an effect that probably can be

attributed to the accumulation of rounding errors in the computer algorithm. In scenario B, the price errors are statistically significant, but small: even for $\Delta = 0.5$, no price error in Table 5B exceeds 0.25 basis points. Interestingly, the bias seems to grow very slowly, or in some cases even decrease, with the caplet maturity. Similar behavior can be seen in the simulation studies by Glasserman and Zhao (1998) for the log-normal market model.

To investigate how the caplet strike impacts the simulation error, Figure 4 graphs some simulated caplet price errors in scenario B for various values of the strike *H*.

Caplet Price Simulation Errors (Simulated - CEV) in Scenario B
1,000,000 Monte Carlo Paths, Log-Euler Scheme, $\Delta = 0.5$, $j(x) = x \cdot MIN(20, x^{-1/2})$

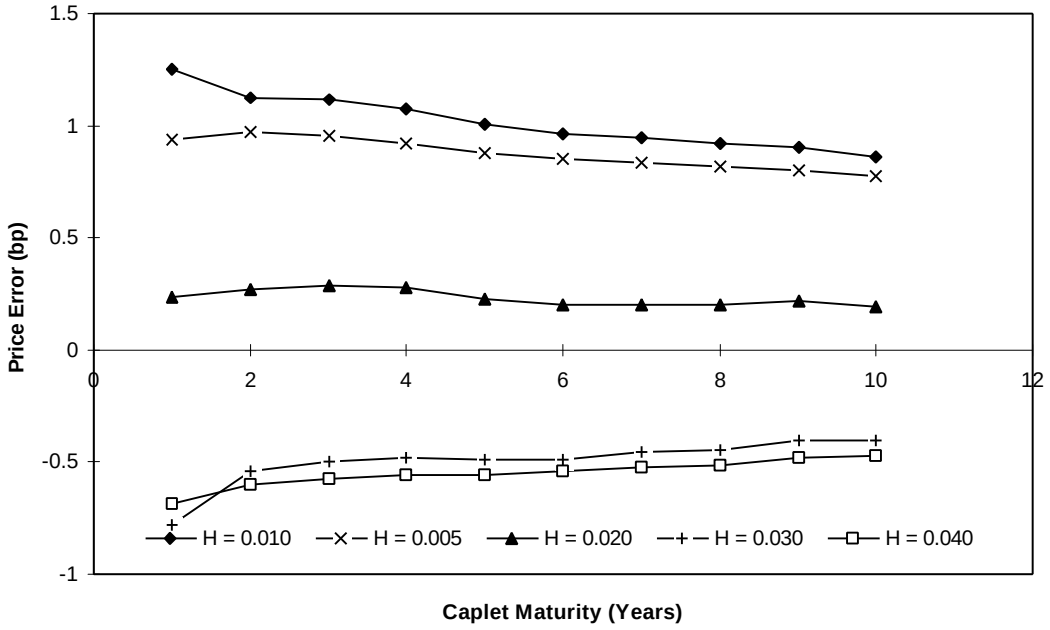

Figure 4

We notice that the errors for out-of-the-money and, in particular, in-the-money caplets are somewhat higher than was the case for at-the-money strikes. In general one expects any approximation of the forward rate distributions to deteriorate in the tails, so these results for the log-Euler scheme are not surprising.

Having investigated the accuracy of the simulation schemes on caps, we now proceed to use Monte Carlo simulation to investigate the accuracy of the swaption pricing formulas in Section 5. As for caps, consider the LCEV model $j(x) = x \cdot MIN(20, x^{-1/2})$. We first turn to the pricing of ATM payer swaption in the scenarios A and B introduced earlier. To keep the simulation bias low, we set the simulation time-step to $\Delta = 0.125$ in the log-Euler discretization scheme (23). The tables below contain simulated swaption prices and sample errors, as well as

prices and price errors (formula price - simulated price) computed by the closed-form approximation in Theorem 5. In the tables, the notation "1 x 5", say, denotes a 1-year swaption on a 5-year swap (that is, the last payment on the swap occurs at year 6).

ATM Payer Swaption Prices in Scenario A (in basis points)
1,000,000 Monte Carlo Paths, $\Delta = 0.125$, Log-Euler Scheme, $j(x) = x \cdot MIN(20, x^{-1/2})$

Swaption	Simulated Price	S.D.	CEV Approx.	Price Error	Price Error (%)
1 x 1	44.13	0.04	44.01	-0.12	-0.27
1 x 5	196.60	0.15	196.18	-0.42	-0.22
1 x 10	342.58	0.25	342.15	-0.43	-0.13
5 x 1	77.44	0.06	77.27	-0.17	-0.22
5 x 5	344.78	0.24	344.45	-0.33	-0.09
5 x 10	599.65	0.37	600.76	1.11	0.18
10 x 1	80.71	0.05	80.76	0.05	0.06
10 x 5	359.17	0.19	360.02	0.86	0.24
10 x 10	623.46	0.24	627.91	4.45	0.71

Table 6A

ATM Payer Swaption Prices and Biases in Scenario B (in basis points)
1,000,000 Monte Carlo Paths, $\Delta = 0.125$, Log-Euler Scheme, $j(x) = x \cdot MIN(20, x^{-1/2})$

Swaption	Simulated Price	S.D.	CEV Approx.	Price Error	Price Error (%)
1 x 1	38.46	0.04	38.25	-0.21	-0.54
1 x 5	70.26	0.06	70.37	0.11	0.16
1 x 10	93.72	0.08	93.90	0.18	0.19
5 x 1	39.73	0.04	39.57	-0.16	-0.40
5 x 5	84.50	0.07	84.62	0.13	0.15
5 x 10	130.45	0.10	130.52	0.08	0.06
10 x 1	37.18	0.04	37.18	0.004	0.01
10 x 5	90.45	0.07	90.79	0.34	0.37
10 x 10	149.98	0.11	150.36	0.38	0.25

Table 6B

The errors reported in Tables 6A-B are somewhat difficult to interpret as they are caused by three separate effects: the bias of the log-Euler scheme; the approximative nature of the formula in

Theorem 5; and the random variation of the Monte Carlo estimate. Nevertheless, with the majority of the errors in Tables 6A-B being less than 0.5 basis points, the formula in Theorem 5 seems to work well, and certainly appears accurate enough for calibration purposes. Not suprisingly, the largest errors are encountered for long-dated options on long-dated swaps; for instance, in scenario A, the price of the 10 x 10 swaption is mispriced by around 4.5 basis points, or 0.71% of the simulated price. Again, it is difficult to determine how much of this error is caused by the imprecision of the Monte Carlo estimate, and how much is caused by the built-in bias of the formula in Theorem 5. Despite the fact that short-term volatilities in scenario B are much higher than those in Scenario A, the strong mean reversion in Scenario B (see Figure 3) ensures that the volatilities of moderate- to long-dated swaps are very similar to those in Scenario A (see Figure 5). As a consequence, the pricing errors associated with Scenario B are roughly comparable to those in scenario A, particularly when the swaption maturity and swap tenor are not very short.

To test how the approximation in Theorem 5 works for in- and out-of-the-money swaptions, we fix the swap tenor to 5 years and consider various swap coupons and swaption maturities. The table below reports our results for scenario A (results for scenario B were very similar and are omitted).

Payer Swaptions on 5-year Swap in Scenario A (in basis points)
1,000,000 Monte Carlo Paths, $\Delta = 0.125$**, Log-Euler Scheme,** $j(x) = x \cdot MIN(20, x^{-1/2})$

	1 x 5		5 x 5		10 x 5	
Coupon (%)	Simulated Price (S.D)	CEV Approx.	Simulated Price (S.D)	CEV Approx.	Simulated Price (S.D)	CEV Approx.
4.5	622.73 (0.07)	623.28	613.46 (0.18)	613.67	540.64 (0.13)	541.67
5.0	452.79 (0.10)	453.29	511.92 (0.21)	511.97	473.81 (0.15)	474.79
5.5	308.88 (0.14)	309.00	422.44 (0.23)	422.29	413.40 (0.17)	414.33
6.0	196.60 (0.15)	196.18	344.78 (0.24)	344.45	359.17 (0.19)	360.02
6.5	116.37 (0.14)	115.53	278.43 (0.24)	277.92	310.78 (0.20)	311.55
7.0	64.01 (0.11)	63.02	222.53 (0.23)	221.88	267.84 (0.20)	268.53
7.5	32.77 (0.08)	31.84	176.10 (0.22)	175.34	229.97 (0.20)	230.57
8.0	15.63 (0.06)	14.92	138.04 (0.20)	137.19	196.74 (0.19)	197.25
8.5	6.97 (0.04)	6.50	107.20 (0.18)	106.33	167.73 (0.18)	168.15

Table 7

Overall, the swaption formula in Theorem 5 does an excellent job, with the vast majority of computed prices in Table 7 being within 1 basis point of the simulated prices (the largest price error in Table 7 is 1.03 basis point and occurs for the 10 x 5 swaption with a coupon of 4.5%).

To get a clearer picture of the volatility skew being generated by the square-root LCEV process, Figure 5 graphs the implied log-normal volatility (computed from the standard Black-Sholes swaption formula, see e.g. Jamshidian (1997)) of the 5 x 5 swaption for various levels of the fixed coupon. The graph also contains data from Scenario B.[7]

Implied Log-Normal Volatility of 5 x 5 Payer Swaption vs. Fixed Coupon
1,000,000 Monte Carlo Paths, $\Delta = 0.125$**, Log-Euler Scheme,** $j(x) = x \cdot MIN(20, x^{-1/2})$

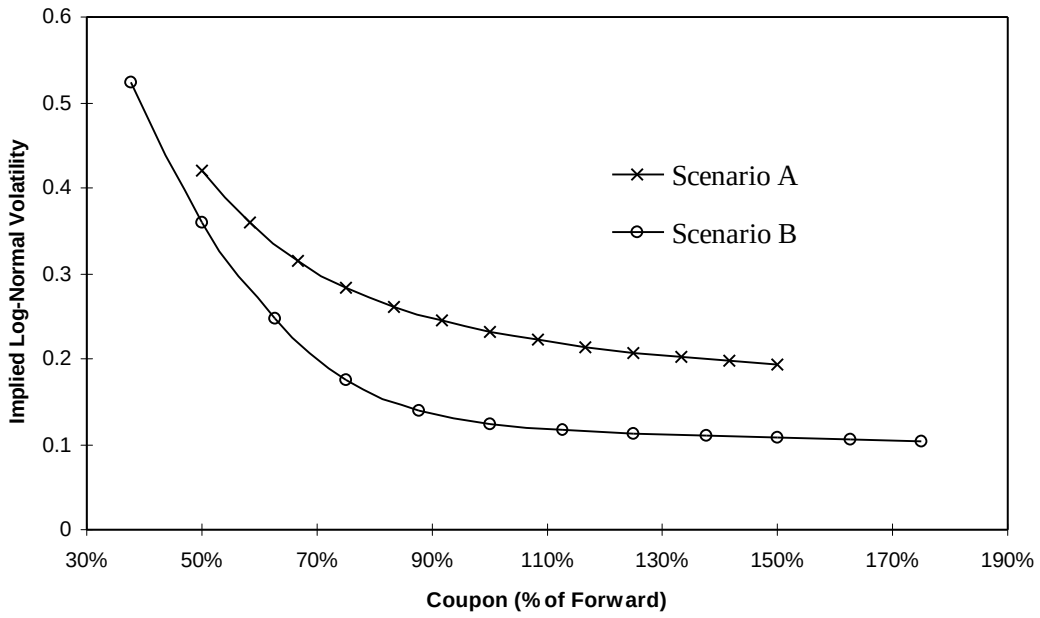

Figure 5

Notice that the volatility skew in Scenario B flattens out for high coupons, a consequence of the downward-sloping volatility function (mean reversion) and the decreasing duration of the fixed side of the swap when coupons are increased.

[7] In the graph, only the simulated swaption prices were used. The prices computed from the formula in Theorem 5 resulted in implied volatilities that were, on average, less than 0.001 away from the ones computed from the simulated prices.

7. Conclusions.

In this paper we have discussed the extension of Libor market models to non-linear forward volatility functions. Unlike log-normal models, the proposed framework is capable of producing volatility smiles or skews consistent with those observed in many markets for caps and swaptions. The effort involved in calibrating the models to market data appears reasonable as efficient numerical routines are available for the pricing of caps and swaptions. Forward rate models based on CEV and LCEV processes were shown to be about as tractable as the log-normal model (which itself belongs to this model class), yet are capable of generating both upward- and downward-sloping volatility skews that conform well to observable market data. For the general case, the paper discusses and demonstrates the usage of Crank-Nicholson finite difference schemes to price caps and swaptions, and contains numerous Monte Carlo studies examining, among other things, the accuracy of the proposed swaption formulas and the price biases introduced by Euler or log-Euler discretizations of the forward rate processes. In general, the proposed model framework seems to be a viable and useful extension of the standard log-normal Libor market models.

References.

Andersen, L. (1995). "Efficient Techniques for Simulation of Interest Rate Models involving Non-Linear Stochastic Differential Equations," Working Paper, General Re Financial Products.

Arnold, L. (1992). *Stochastic Differential Equations: Theory and practice*, Reprint, Krieger Publishing Company.

Black, F. (1976). "The Pricing of Commodity Contracts," *Journal of Financial Economics*, 3, 167-179.

Borodin, A. and P. Salminen (1996). *Handbook of Brownian Motion -- Facts and Formulae*, Birkhaeuser, Basel.

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

Brace, A. (1996). "Dual Swap and Swaption Formulae in the Normal and Lognormal Models," Working Paper, UNSW Australia.

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

Brotherton-Ratcliffe, R. (1997). "The BGM model for Path-Dependent Swaps," Working Paper, General Re Financial Products.

Cox, J. and S. Ross (1976). "The Valuation of Options for Alternative Stochastic Processes," *Journal of Financial Economics*, 3 (January-March), 145-166.

Cox, J., J. Ingersoll, and S. Ross (1985). "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385-407.

Ding, C. G. (1992). "Algorithm AS275: Computing the Non-Central c^2 Distribution function," *Applied Statistics*, 41, 478-482.

Duffie, D. (1996). *Dynamic Asset Pricing Theory*, 2nd ed., Princeton Univ. Press, New Jersey.

Dupire, B. (1994). "Pricing with a Smile," *RISK*, (January), 18-20.

Glasserman, P. and Z. Zhao (1998). "Arbitrage-Free Discretization of Lognormal Forward Libor and Swap Rate Models," Working Paper, Columbia University.

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.

Jamshidian, F. (1989). "An Exact Bond Option Pricing Formula," *Journal of Finance*, 44 (March), 205-209.

Jamshidian, F. (1991). "Forward Induction and Construction of Yield Curve Diffusion Models," Journal of Fixed Income, June, 62-74.

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

Johnson, N. L. and S. Kotz (1981). *Continuous Univariate Distributions, vol. 2*, Wiley and Sons.

Karatzas, I. And S. Shreve (1991). *Brownian Motion and Stochastic Calculus*, Springer Verlag, New York.

Kloeden, P. and E. Platen (1992). *Numerical Solution of Stochastic Differential Equations*, Springer Verlag, New York.

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

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

Schroder, M. (1989). "Computing the Constant Elasticity of Variance Option Pricing Formula," *Journal of Finance*, 44 (1), 211-218.

Sidenius, J. (1997). "Multi-Factor 'Market' Models", Skandinaviska Enskilda Banken, Copenhagen.

Smith, G. D. (1985*). Numerical Solution of Partial Differential Equations: Finite Difference Methods*, Oxford University Press.

Talay, D. and L. Tubaro (1990). "Expansion of the Global Error for Numerical Schemes Solving Stochastic Differential Equations," *Stoch. Analysis and Applications*, 8, 94-120.

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

Appendix A-- Proofs of Lemmas and Theorems

Proof of Theorem 1.

Due to the recursive nature of the SDEs in the various measures, it suffices to consider the system of SDEs (2) under the spot measure **Q**:

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

Assume that existence and uniqueness has been established for F_j, $j = n(t)..k-1$. Also assume that $F_j \geq 0$, $j = n(t),..,k$. By elementary analysis, the local Lipschitz and growth conditions on j ensures that each term in the sum for $m_k(t)$ is locally Lipschitz continuous and bounded. The growth condition on j in turn ensures that the product $j\left(F_k(t)\right) l_k^{\mathrm{T}}(t) m_k(t)$ is also locally Lipschitz continuous and, due to the boundedness of $m_k(t)$, satisfies a linear growth condition. For $F_k \geq 0$, the result in Theorem 1 now follows from standard existence and uniqueness theorems (e.g. Arnold (1992), p. 112). Clearly, if $F_k(t) = 0$ for some $t < \infty$, uniqueness and $j(0) = 0$ ensures that $F_k(s) = 0$ for all $s > t$, i.e. the forward rate is non-negative.

To show that forward rates stay strictly positive if started above 0, consider first the probability measure $\mathbf{Q}^{k+1}$. According to Lemma 1, under $\mathbf{Q}^{k+1}$ the boundary behavior of F_k in 0 is the same as that of the process *x*,

$$dx(v) = j\left(x(v)\right)dZ(v) \qquad \text{(A.1)}$$

where $x(0) > 0$ and $Z(t)$ is a one-dimensional Brownian motion. According to Borodin and Salminen (1996, p.17), *x* will have 0 as an inaccessible boundary if

$$\int_0^z m((a,z))s(da) = +\infty,$$

where

$$m(dx) = 2j(x)^{-2}dx, \quad s(dx) = dx,$$

are the speed and the scale measures of the process (A.1), respectively. The local Lipshitz condition together with $j(0) = 0$ gives us that for $b \in (a,z)$:

$$|j(b)| = |j(b) - j(0)| \le K_z b \quad \Rightarrow \quad j(b)^{-2} \ge K_z^{-2} b^{-2}$$

and thereby,

$$\int_0^z m((a,z))s(da) = \int_0^z \int_a^z 2j(b)^{-2} db\, da \ge \int_0^z \int_a^z 2K_z^{-2} \cdot b^{-2} dbda = +\infty.$$

Hence we conclude that F_k cannot reach 0 under $\mathbf{Q}^{k+1}$. To extend the result to other probability measures, notice that since the terms of the form $j(x)/(1+ d_j x)$ are bounded for non-negative *x*, the *Novikov condition* (Duffie (1996, p. 288)) guarantees that $\mathbf{Q}^i$, $\mathbf{Q}^j$ are equivalent probability measures for all *i*, *j*. As equivalent measures share null-sets, $\mathbf{Q}^{k+1}(F_k(t) = 0) = 0$ for all *t* shows that F_k stays positive under all probability measures.♠

Proof of Lemma 2.

a) follows from the remark in Karatzas and Shreve (1991, p. 332); and b) follows from example 5.2.15 in Karatzas and Shreve (1991). As in the proof of Theorem 1, we characterize the boundary condition at 0 through a speed/scale measure integral. In particular, 0 is an accessible boundary for (7) iff

$$\int_0^z m((a,z))s(da) < \infty,$$

where

$$m(dx) = 2x^{-2a}dx, \; s(dx) = dx.$$

Hence,

$$\int_0^z m((a,z))s(da) = \int_0^z \int_a^z m(db)\, s(da) = \int_0^z \int_a^z 2b^{-2a}\, db\, da \begin{cases} < +\infty & , a < 1 \\ = +\infty & , a \geq 1 \end{cases}$$

which shows (c). To show (d), according to Borodin and Salminen (1996, Chapter II) it suffices to consider the integral

$$\int_0^z \big(s(z) - s(a)\big) m(da)$$

If the integral is bounded by $+\infty$, the level $x = 0$ is a so-called *entrance* point. If $x = 0$ is also an attainable boundary, the SDE (7) needs to be associated with additional boundary conditions in $x = 0$ for it to have a unique solution. For our case we get

$$\int_0^z \big(s(z) - s(a)\big) m(da) = 2\int_0^z (z-a) a^{-2a}\, da \begin{cases} < +\infty & , a < 1/2 \\ = +\infty & , a \geq 1/2 \end{cases}$$

Hence, for the case of $0 < a < 1/2$ additional boundary conditions have to be added for the process (7) to have a unique solution. ♠

Proof of Lemma 3.
Using Lemma 1, we see that it suffices to consider the time-changed forward rate process

$$df_k(v) = f_k(v)^a\, d\tilde{Z}_{k+1}^k(v)$$

where $\tilde{Z}_{k+1}^k$ is a one-dimensional Brownian motion. Assuming that $X_k(t) > 0$, we define $x_k\big(v_k(s)\big) = X_k(s)$ which, according to Ito's Lemma, satisfies

$$dx_k(v) = \frac{1-2a}{1-a}\, dv + 2\sqrt{x_k(v)}\, d\tilde{Z}_{k+1}^k(v)$$

up to the stopping time $\inf_{s>t}\{x_k(s) = 0\}$. The SDE for $x_k(v)$ identifies it as a squared Bessel process of index $J = -2/(1-a)$. Using a result in Borodin and Salminen (1996, p. 117), we obtain the results of Lemma 3 after substituting $v \equiv v_k(t,T)$. ♠

Proof of Theorem 2.

Let $E_t^{k+1}[\cdot]$ denote expectation under $\mathbf{Q}^{k+1}$ conditional on the information available at time *t*. Absence of arbitrage implies that

$$C_k(t) = d_k P(t,T_{k+1}) E_t^{k+1}\left[(F_k(T_k)-H)^+\right] = d_k P(t,T_{k+1}) E_{v_k(t)}^{k+1}\left[\left(f_k(v_k(T_k))-H\right)^+\right],$$

where the notation from (5) is used in the second equality. (5) and the Feynman-Kac Theorem (Duffie (1996)) proves Theorem 2 after a suitable change of time-variable.♠

Proof of Theorem 3.

For the case a), the results of Lemma 3 enable us to write

$$E^{k+1}\left[\left(F_k(T_{k+1})-H\right)^+ | F_k(t)\right] = \int_{a/2}^{+\infty}\left(\left(2(1-a)vx\right)^{\frac{1}{2(1-a)}} - H\right) y\ (x)dx,$$

where

$$x = \frac{F_k(T_k)^{2(1-a)}}{2(1-a)^2 v(t,T_k)},\ x_0 = \frac{F_k(t)^{2(1-a)}}{2(1-a)^2 v(t,T_k)},\ y\ (x) = e^{-x-x_0}\left(\frac{x_0}{x}\right)^{\frac{1}{4(1-a)}} I_{\frac{1}{2(1-a)}}\left(2\sqrt{x_0 x}\right).$$

We now note that

$$\begin{aligned}\left(2(1-a)^2 v(t,T_k)\right)^{\frac{1}{2(1-a)}} \int_{a/2}^{+\infty} x^{\frac{1}{2(1-a)}} y\ (x)dx &= F_k(0)\int_{a/2}^{+\infty} e^{-x-x_0}\left(\frac{x}{x_0}\right)^{\frac{1}{4(1-a)}} I_{\frac{1}{2(1-a)}}\left(2\sqrt{x_0 x}\right)dx \\ &= F_k(0)\int_a^{+\infty} e^{-\frac{y-2x_0}{2}}\left(\frac{y}{2x_0}\right)^{\frac{1}{4(1-a)}} I_{\frac{1}{2(1-a)}}\left(\sqrt{2x_0 y}\right)dy\ .\end{aligned}$$

Comparing to Johnson and Kotz (1981, p. 436), we can now identify the first part of formula in a). To derive the second part, let $p(x, J, l) = \partial c^2(x, J, l)/\partial x$. Using the results of Johnson and Kotz (1981, p. 436) we note that $y\ (x) = p(2x_0, 2+1/(1-a), x)$. Schroder (1989) shows that

$$\int_0^{z/2} p(2x, J, 2l)dl = 1 - c^2(2x, J-2, z).$$

Using this we obtain

$$\int_{a/2}^{+\infty} y\ (x)dx = c^2(c,b,a),$$

which shows a). The result c) can be derived in the same way.♠

Proof of Theorem 4.

We first concentrate on proving convergence in distribution, i.e. that $\lim\left|\mathbf{P}(x(T)<h)-\mathbf{P}(y(T)<h)\right|=0$. The result in a) is obvious and follows directly from the non-explosiveness of the CEV process (see Lemma 2) which implies that the probability of *y* hitting e, and thus of *y* differing from *x*, vanishes for $e\to\infty$. This argument obviously cannot be used to prove b) since, as we have already seen, CEV processes with $0<a<1$ have positive probability of hitting 0. So define the stopping time

$$t = \inf\{v\colon x(v)\le e\}$$

so that $x(v)=y(v)$ for $0\le v\le t$. We can then write

$$\begin{aligned}\left|\mathbf{P}(x(T)<h)-\mathbf{P}(y(T)<h)\right| &= \left|\mathbf{P}(x(T)<h,\ t<T)-\mathbf{P}(y(T)<h,\ t<T)\right| \\ &= \left|\mathbf{P}(y(T)\ge h,\ t<T)-\mathbf{P}(x(T)\ge h,\ t<T)\right|.\end{aligned} \tag{A.2}$$

To bound[8] the terms in (A.2), we first introduce the indicator function $1_{\{t<T\}}$ and note that

$$E_{\min(t,T)}\left[1_{\{t<T\}}\,y(T)\right]=1_{\{t<T\}}\,E_{\min(t,T)}\left[y(T)\right]=1_{\{t<T\}}\,y(\min(t,T))\le e, \tag{A.3}$$

where the second equality follows from the Optional Sampling Theorem (Karatzas and Shreve (1991), p.19) and the fact that *y* is a martingale. By the law of iterated conditional expectations, we conclude that

$$E_0\left[1_{\{t<T\}}\,y(T)\right]\le e.$$

Also note that

$$\mathbf{P}(y(T)\ge h,\ t<T)=\int_{\{y(T)\ge h,\ t<T\}} d\mathbf{P}\le\int_{\{y(T)\ge h,\ t<T\}}\frac{y(T)}{h}\,d\mathbf{P}\le\int_{\{t<T\}}\frac{y(T)}{h}\,d\mathbf{P}\le\frac{e}{h}, \tag{A.4}$$

[8] We are grateful to Steven Shreve for pointing out this ingenious argument.

where the last inequality follows from (A.3). With the boundary conditions in 0 as specified in the Theorem, the process *x* is also a martingale (see discussion after Lemma 2), and we can use the same arguments to derive

$$\mathbf{P}\big(x(T) \geq h,\ t < T\big) \leq \frac{e}{h} \tag{A.5}$$

Now, finally,

$$\big|\mathbf{P}\big(x(T) < h\big) - \mathbf{P}\big(y(T) < h\big)\big| \leq MAX\big(\mathbf{P}\big(y(T) \geq h,\ t < T\big),\ \mathbf{P}\big(x(T) \geq h,\ t < T\big)\big) \leq \frac{e}{h},$$

where the second inequality follows from (A.4) and (A.5). Letting $e \to 0_+$ gives the result in b).

Having established convergence in distribution, we now need to show convergence of $E^{\mathbf{P}}\big[(x(T)-H)^+\big]$ - $E^{\mathbf{P}}\big[(y(T)-H)^+\big]$. We have

$$\begin{aligned} E^{\mathbf{P}}\big[(x(T)-H)^+\big] - E^{\mathbf{P}}\big[(y(T)-H)^+\big] &= E^{\mathbf{P}}\big[x(T) + (H-x(T))^+\big] - E^{\mathbf{P}}\big[y(T) + (H-y(T))^+\big] \\ &= E^{\mathbf{P}}\big[(H-x(T))^+\big] - E^{\mathbf{P}}\big[(H-y(T))^+\big] \end{aligned}$$

where the second inequality follows from the martingale property of *x* and *y* and the fact that $x(0) = y(0)$. As $(H-z)^+$ is a bounded, continuous function of *z*, the results in the Theorem follow from the Dominated Convergence Theorem (see Duffie (1996), p. 280).♠

Appendix B -- Finite difference solution of (9)

As a first step, introduce a uniform mesh (x_l, t_j) with

$$x_l = x_0 + l\Delta_x,\ \ t_j = j\Delta_t$$

for $0 \leq l \leq N$, $0 \leq j \leq M$. x_0 and x_N represent the upper and lower limits of *x*-space (typically either 0 and ∞, or $-\infty$ and ∞), and should be set such that most of the statistically significant *x*-space is captured by the mesh. t_M represents the largest value of t for which a solution of (9) will be required. Using $\hat{g}_{l,j}$ to denote the finite difference approximation of the true solution $g(t_l, x_j)$, a Crank-Nicholson discretization of (9) at node (x_l, t_j) is given by

$$-\frac{1}{\Delta_t}(\hat{g}_{l,j+1}-\hat{g}_{l,j})+\frac{j\,(x_l)^2}{2\Delta_x^2}\left(\hat{g}_{l+1,j}-2\hat{g}_{l,j}+\hat{g}_{l-1,j}+\hat{g}_{l+1,j+1}-2\hat{g}_{l,j+1}+\hat{g}_{l-1,j+1}\right)=0$$

or, equivalently,

$$(1+a_l)\hat{g}_{l,j+1}-\tfrac{1}{2}a_l\hat{g}_{l+1,j+1}-\tfrac{1}{2}a_l\hat{g}_{l-1,j+1}=(1-a_l)\hat{g}_{l,j}+\tfrac{1}{2}a_l\hat{g}_{l+1,j}+\tfrac{1}{2}a_l\hat{g}_{l-1,j}\,, \quad \text{(B.1)}$$

where $a_l = j\,(x_l)^2\Delta_t/\Delta_x^2$. Starting from the known boundary condition at $t=0$, $\hat{g}_{i,0}=g(0,x_i)=(x_i-H)^+$, the tridiagonal system (B.1) can be solved forward in t using, for example, the *O*(*N*) algorithm in Press *et al* (1992), p. 51.

The close resemblance between (6) and the heat equation strongly suggests that (B.1) is stable. This can be verified formally by a Von Neuman analysis (Smith (1985)). A Taylor-expansion analysis of (B.1) further reveals that the local truncation error of (B.1) approaches 0 as Δ_t and Δ_x approach 0 ("consistency"). By *Lax's equivalence theorem* for linear PDEs (e.g. Smith (1985), p. 72) we conclude that (B.1) is *convergent*.