

Yield Curve Modelling with Skews and Stochastic Volatility

by

Leif Andersen and Jesper Andreasen

Bank of America Securities

April 2002

Modified August 2002

Abstract

This paper discusses a variety of techniques for modeling the evolution of interest rates in the presence of stochastic volatility. We cover both Libor Market models and low-dimensional Markovian HJM models. To facilitate model calibration, special emphasis is put on efficient pricing of plain-vanilla instruments. Effects of stochastic volatility on CMS structures as well as Bermudan swaptions are discussed.

1. Introduction.

There is currently great interest in improving fixed income models to better capture observed volatility skews and smiles in cap and swaption markets. Approaches suggested in the literature include state-dependent diffusions (see e.g. Andersen and Andreasen, 2000) and geometric Brownian motion overlaid with a jump process (as in Glasserman and Kou, 1999). While these approaches are both reasonable, they are not without problems and it is unlikely that they tell the whole story. For instance, as discussed in Rebonato (2001) and in a number of empirical studies, inclusion of stochastic volatility into fixed income models can significantly improve the realism of the models. This paper continues this line of research by focusing on flexible, yet practical techniques for the financial engineer to construct interest rate models with stochastic volatility. We work with both multi-factor Libor Market (LM) models and low-dimensional Markovian HJM models, and demonstrate that the proposed models are reasonable from an empirical perspective. To aid calibration and fast mark-to-market of simple derivatives, we pay particular attention to the development a variety of exact and approximative techniques to compute the prices of basic European derivatives such as caps, swaptions, and CMS options. Further, the paper introduces numerical methods for complex derivatives and gives examples of the effects of stochastic volatility on Bermudan swaptions.

2. Notation. Libor Market model.

Let $\{T_i\}_{i=1}^N$ be a discrete tenor structure and let $F_k(t)$ denote the time t value of the discretely compounded forward rate spanning $[T_k, T_{k+1}]$. Also, let $R_{s,e}(t)$ denote the par rate of a plain-vanilla swap exchanging fixed for floating payments at points in the tenor structure $T_{s+1}, T_{s+2}, \dots, T_e$. With $\mathbf{d}_i \equiv T_{i+1} - T_i$ and $P(t, T)$ being the time t discount factor to time T , we have

$$F_k(t) = \mathbf{d}_k^{-1} (P(t, T_k) / P(t, T_{k+1}) - 1), \quad t \leq T_k;$$

$$R_{s,e}(t) = \frac{P(t, T_e) - P(t, T_s)}{A_{s,e}(t)}, \quad A_{s,e}(t) = \sum_{i=s+1}^e \mathbf{d}_{i-1} P(t, T_i), \quad t \leq T_s.$$

We now postulate model arbitrage-free forward rate dynamics of the form

$$dF_k(t) = \mathbf{j}(F_k(t)) \sqrt{V(t)} \mathbf{I}_k(t)^T \left[\sqrt{V(t)} \mathbf{m}_k(t) dt + dW(t) \right] \quad (1a)$$

where W is an n -dimensional Brownian motion, $\mathbf{j} : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is a well-behaved deterministic function satisfying $\mathbf{j}(0) = 0$, \mathbf{I}_k is an n -dimensional deterministic volatility function loading each scalar Brownian motion, $V(t)$ is a scalar positive process to be specified, and \mathbf{m}_k is an n -dimensional numeraire-specific drift that ensures lack of arbitrage across bonds. An expression for \mathbf{m}_k can be found in Andersen and Andreasen (2000); for the special case where the bond $P(t, T_{k+1})$ is used as the numeraire asset $\mathbf{m}_k(t) = 0$ and the k th forward rate is a martingale. Notice that (1a) is non-standard as we have allowed the local variance $\|\mathbf{I}_k(t)\|^2$ to be multiplied by a scalar factor $V(t)$. Setting $V(t) \equiv 1$ (or some

deterministic function) recovers the extended Libor market model of Andersen and Andreasen (2000); letting V be random introduces the desired stochastic movements of rate volatilities. We note that these movements of the volatility surface are "parallel"; additional factors could, in principle, be added to make the fluctuations of volatilities more complicated, but the resulting abundance of hard-to-estimate process parameters would make the model difficult to populate, hard for traders to comprehend, and probably not much more flexible in terms of the types of volatility smiles and smirks that could be generated. The process for V is here assumed to be a classical mean-reverting process:

$$dV(t) = \mathbf{k}(\mathbf{q} - V(t))dt + \mathbf{e}\mathbf{y}(V(t))dZ(t), \quad (1b)$$

where $\mathbf{k}, \mathbf{q}, \mathbf{e} > 0$, $\mathbf{y} : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is a smooth function with $\mathbf{y}(0) = 0$, and Z is a scalar Brownian motion independent of W . V has the role of a scale factor, so typically (but not necessarily) $V(0) = \mathbf{q} = 1$. To ensure proper scale behavior, it is most natural to set $\mathbf{y}(x) = x^p$, $p > 0$, but for now we do not explicitly impose this choice. A few comments about the chosen framework:

- The volatility functions $I_k, k = 1, 2, \dots, N-1$ are assumed exogenous and it is up to the model builder to decide whether a parametric approach or a non-parametric approach should be used in estimating this function.
- The assumptions of $\mathbf{y}(0) = 0$ and $\mathbf{j}(0) = 0$ ensure that variances and forward rates cannot go below zero. We shall occasionally violate the latter condition to gain tractability.
- The V -process generates a near-symmetric smile which is superimposed onto the base smile produced by the function \mathbf{j} . The smile generated by the V -process is "stationary" in the sense that the bottom of the smile will move along with fluctuations in the forward rate.
- If $\mathbf{j}(x)/x$ is a monotonically downward sloping function in x (e.g. $\mathbf{j}(x) = x^p, 0 < p < 1$), the base smile is a downward-sloping skew. If $\mathbf{j}(x)/x$ is non-monotonic (e.g. $\mathbf{j}(x) = x^p + wx^q, 0 < p < 1, q > 1$), the base smile can be a true U -shaped smile, but will be non-stationary: when rates move, the bottom of the smile will not move with the forward.
- The smile effects of \mathbf{j} typically persist for much longer maturities than those of the stochastic volatility process (1b), a consequence of the fact that \mathbf{j} introduces dependency in log-increments of forward rate movements. Working with a general skew function \mathbf{j} thus allows the model builder some flexibility in shaping the long-term behavior of the smile (but see the point above for a caveat).
- While Z is assumed independent¹ of W , as long as $\mathbf{j}(x)/x \neq \text{const.}$ our model nevertheless is able to generate a range of effective correlations between the local forward rate volatility (defined as $\mathbf{s}_k(t, F_k, V) = \sqrt{V} \sqrt{\|I_k(t)\|^2 \mathbf{j}(F_k)/F_k}$) and the forward itself.

¹ The need to assume independence is primarily technical: without it, any change of probability measure introduces terms depending on forward rates in the process for V , thereby destroying the tractability of the model.

As shown in Andersen and Brotherton-Ratcliffe (2001), (1a-b) leads to swap par rate dynamics that for many practical choices of \mathbf{j} can be closely approximated by

$$dR_{s,e}(t) \approx \sqrt{V(t)} \mathbf{j} (R_{s,e}(t)) \mathbf{I}_{s,e}(t) dW_{s,e}(t) \quad (2)$$

where $W_{s,e}$ is a vector Brownian motion under the measure (“swap measure”) induced by using the annuity factor $A_{s,e}$ as numeraire, and where $\mathbf{I}_{s,e}(t)$ as usual can be approximated as a linear combination of the \mathbf{I}_k 's, $k = s, s+1, \dots, e-1$. Applications and tests of swap rate approximations such as (2) can be found in numerous sources; see for instance Andersen and Andreasen (2000), Glasserman and Kou (1999), and Rebonato (2001), to name a few.

3. Cap and Swaption Pricing by Asymptotic Expansions.

To enable fast model calibration, we now seek efficient means of computing prices of European caps and swaptions. Specifically, consider now a caplet C_k paying at time T_{k+1} the amount $C_k(T_{k+1}) = \mathbf{d}_k (F_k(T_k) - X)^+$, as well as a European payer swaption $S_{s,e}$ paying at time T_s the amount $S_{s,e}(T_s) = A_{s,e}(T_s)(R_{s,e}(T_s) - K)^+$ (X and K are thus the caplet strike and the fixed swap coupon, respectively). We have

$$\begin{aligned} C_k(0) &= \mathbf{d}_k P(0, T_{k+1}) E^{k+1} (F_k(T_k) - X)^+, \\ S_{s,e}(0) &= A_{s,e}(0) E^{A_{s,e}} (R_{s,e}(T_s) - K)^+, \end{aligned}$$

where E^{k+1} and $E^{A_{s,e}}$ denote expectations with respect to the probability measures induced by the numeraires $P(t, T_{k+1})$ and $A_{s,e}(t)$, respectively. F_k is a martingale under the former measure; $R_{s,e}$ is a martingale under the latter. Notice that with the processes (1a) and (2) having identical form, caplet and swaption price formulas will be completely equivalent (short of numeraire scaling factors), and in this section we only consider the former. Specifically, we here propose an asymptotic expansion for the general caplet pricing problem; in the next section we present an exact transform-based formula for special choices of the function \mathbf{j} and \mathbf{y} .

To develop an asymptotic expansion, consider first the special case where V is non-random, e.g. $V \equiv 1$. Even for this simple case, caplet prices cannot be computed explicitly under our process assumption. However, for a large class of \mathbf{j} 's, accurate asymptotic expansions can be constructed. For instance, a small-time expansion around the special case $\mathbf{j}(x) = x$ results in the following expression:

Caplet Expansion for $V \equiv 1$.

Assume $V(t) = 1$ for all t . Defining $c = T_k^{-1} \int_0^{T_k} \|\mathbf{I}_k(t)\|^2 dt$ and writing $C(0) = P(0, T_{k+1}) \mathbf{d}_k g(F_k(0); c)$ for some function $g = g(F; c)$, we have

$$g(F; c) = F \Phi(d_+) - X \Phi(d_-), \quad d_{\pm} = \frac{\ln(F/X) \pm \frac{1}{2} \Omega(F, c)^2}{\Omega(F, c)}, \quad (3)$$

where Φ is the cumulative Gaussian distribution function, and

$$\Omega(F, c) = \Omega_0(F)c^{1/2}T_k^{1/2} + \Omega_1(F)c^{3/2}T_k^{3/2} + O(T_k^{5/2}),$$

$$\Omega_0(F) = \frac{\ln(F/X)}{\int_X^F \mathbf{j}(u)^{-1} du}; \quad \Omega_1(F) = - \frac{\Omega_0(F)}{\left[\int_X^F \mathbf{j}(u)^{-1} du \right]^2} \ln \left[\frac{\Omega_0(F)}{\int_X^F \mathbf{j}(u)^{-1} du} \right] \frac{FX}{\left[\int_X^F \mathbf{j}(u)^{-1} du \right]^{1/2}}.$$

Proof and tests of this result can be found in Andersen and Brotherton-Ratcliffe (2001) who also list the result for the limit $F_k \rightarrow X$. While being constructed to be accurate for small maturities, the expansion turns out to often retain its precision for long maturities, even for options with strikes far from at-the-money. This is, for instance, the case for CEV-like specifications $\mathbf{j}(x) = x^p$, as well as truncated exponentials, $\mathbf{j}(x) = x(1 + ae^{-bx})$, $a, b > 0$.

We now relax the assumption that $V \equiv 1$ and instead assume that V follows (1b). A small-time expansion for the resulting caplet pricing equations turns out to be cumbersome; instead we adapt the small- ϵ expansions in Hull and White (1987) and Lewis (2000) to our purposes. The result is:

Caplet Expansion for Stochastic V .

Define $Y = \ln(F_k(0)/X)$ and $\bar{c}(V) = T_k^{-1} \int_0^{T_k} \|\mathbf{I}_k(t)\|^2 (\mathbf{q} + [V - \mathbf{q}]e^{-kt}) dt$. Then $C(0) = P(0, T_{k+1})(T_{k+1} - T_k)g(F_k(0); c^*(V(0)))$ with g given as in Eq. (3), and

$$c^*(V) = \bar{c}(V) + (\mathbf{a}_0 \mathbf{e}^2 + \mathbf{b}_0 \mathbf{e}^4) + (\mathbf{a}_1 \mathbf{e}^2 + \mathbf{b}_1 \mathbf{e}^4)Y^2 + \mathbf{b}_2 \mathbf{e}^4 Y^4 e^{-\Lambda \mathbf{e}^2 Y^2} + O(\mathbf{e}^6) \quad (4)$$

for coefficients $\mathbf{a}_0, \mathbf{a}_1, \mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2$, and Λ an arbitrary positive number. The corresponding Black-Scholes implied volatility is given by

$$\mathbf{s}_{imp} = \Omega_0 \sqrt{c^*(V)} + \Omega_1 c^*(V)^{3/2} T_k + O(T_k^2).$$

The coefficients $\mathbf{a}_0, \mathbf{a}_1, \mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2$ depend on the parameters of the process of V , including the chosen skew function \mathbf{y} . Their computation is essentially a matter of tedious algebra, but the resulting expressions are lengthy and for space reasons we here must refer to Andersen and Brotherton-Ratcliffe (2001) for details (article is available on the Internet). Suffice to say that all coefficients can be computed almost instantaneously on a computer, making the resulting expression extremely efficient and fast enough even for production systems pricing and risk-managing many thousands of caplets. The parameter Λ is, essentially, a defense mechanism protecting the expansion from possible degeneration for options very deeply out-of-the money (where $|Y|$ can grow large). In most cases, the exact value of Λ is of little importance.

We note that for many typical parameters and option contracts, it often suffices to only include terms in (4) to order $O(\mathbf{e}^2)$ (that is, set $\mathbf{b}_0 = \mathbf{b}_1 = \mathbf{b}_2 = 0$), but the higher-order terms often become important when, say, the mean reversion speed (\mathbf{k}) is low and the volatility-of-variance (\mathbf{e}) is high. Figure 1 demonstrates typical performance of the $O(\mathbf{e}^2)$ and $O(\mathbf{e}^4)$ expansions (“O4” and “O2” in the figure). Many further tests can be found in Andersen and Brotherton-Ratcliffe (2001).

4. Exact Transform Solution.

In cases where very high volatility of variance and/or very low mean reversions are required, the expansions in Section 3 might have to be continued to inconveniently high order. In such situations it is sometimes safer and easier to instead rely on exact transform-based pricing expressions. To develop such expressions, however, we must simplify our setup somewhat. In particular, we specialize to an affine setting with $\mathbf{y}(x) = \sqrt{x}$ and $\mathbf{j}(x) = mx + (1-m)L$, where $L \geq 0$ and $0 < m \leq 1$. As we shall show, this setup is tractable, although the tractability comes with the cost of permitting interest rates can become negative with positive probability. We also notice that the choice of $\mathbf{y}(x) = \sqrt{x}$ allows the process (2) to reach 0 with positive probability if, as will often be the case in applications, $2\mathbf{k}\mathbf{q} < \mathbf{e}^2$. As 0 is non-absorbing, this has few practical ramifications, however.

As demonstrated in Section 3, caplet and swaption formulas have the same form in our setup; for variety we will here develop the latter. Dropping subscripts (i.e. $R_{s,e}(t)$ becomes $R(t)$), and so forth, we write

$$S(0) = A(0)E^A(R(T_s) - K)^+ = \frac{A(0)}{m}E^A(\mathbf{x}(T_s) - K')^+ \equiv \frac{A(0)}{m}f(0)$$

where $K' = mK + (1-m)L$ and

$$\frac{d\mathbf{x}(t)}{\mathbf{x}(t)} = \sqrt{V(t)}\mathbf{I}(t)m dW(t), \quad \mathbf{x}(0) = mR(0) + (1-m)L.$$

Using the results of Heston (1993) and Lewis (2000) we find the swaption price formula by numerically solving the inverse Fourier integral

$$f(0) = \mathbf{x}(0) - \frac{K'}{2\mathbf{p}} \int_{-\infty}^{+\infty} \frac{e^{(-i\mathbf{w} + \frac{1}{2})\ln \mathbf{x}(0)/K'}}{\mathbf{w}^2 + \frac{1}{4}} H(0, \mathbf{w}) d\mathbf{w} \quad (5)$$

where $i = \sqrt{-1}$ and $H(t, \mathbf{w}) = e^{a(t, \mathbf{w}) + b(t, \mathbf{w})V(t)}$ with a, b given as the solutions to the Riccati ODEs

$$\begin{aligned} \frac{da}{dt} &= -\mathbf{k}\mathbf{q}b, \quad a(T_s; \mathbf{w}) = 0, \\ \frac{db}{dt} &= \frac{1}{2}m^2\mathbf{I}(t)^2\mathbf{w}^2 + \mathbf{k}b - \frac{1}{2}\mathbf{e}^2b^2, \quad b(T_s; \mathbf{w}) = 0. \end{aligned}$$

For constant \mathbf{I} these ODEs can be solved in closed form (see e.g. Lewis (2000) or Lipton (2001)), a result which can also be used iteratively for piecewise constant \mathbf{I} . In the general case, a simple numerical Runge-Kutta scheme can be applied at little additional computational expense. We point out that the transform solution can easily accommodate non-zero correlation \mathbf{r} between V and R by modifying the term $\mathbf{k}b$ in the second ODE to $b(\mathbf{k} + (i\mathbf{w} - \frac{1}{2})\mathbf{e}\mathbf{I}\mathbf{r})$, but we shall rarely need this as non-zero correlation will hamper the construction of practical models for the joint dynamics of the full yield curve (see footnote 1).

We finally note that when computing the integral (5) numerically, performance can often be improved by using the case $\mathbf{e} = 0$ as a ‘‘control variate’’ through the split

$$f(0) = \mathbf{x}(0) - \frac{K'}{2p} \int_{-\infty}^{+\infty} \frac{e^{(-i\mathbf{w} + \frac{1}{2})\ln \mathbf{x}(0)/K'}}{\mathbf{w}^2 + \frac{1}{4}} \left(H(0, \mathbf{w}) - e^{-(\mathbf{w}^2 + \frac{1}{4})v^2/2} \right) d\mathbf{w} - K' g_{BS}$$

where

$$g_{BS} = \Phi(x_-) + \frac{\mathbf{x}(0)}{K'} (1 - \Phi(x_+));$$

$$x_{\pm} = \frac{\ln \mathbf{x}(0)/K'}{v} \pm \frac{1}{2}v, \quad v^2 = \int_0^{T_s} m^2 \mathbf{I}(u)^2 (\mathbf{q} + (V(0) - \mathbf{q})e^{-ku}) du.$$

With this trick, the combined scheme of Runge-Kutta and numerical Fourier inversion is fast and accurate, although some attention to step sizes is, as always, needed for long-dated options with strikes far from the at-the-money point. Typically, accurate option prices can be obtained in around 0.03 seconds per option on a regular PC (about three times faster for the constant parameter case).

5. A low-dimensional Markov model with stochastic volatility.

In Section 2, we specified our stochastic interest rate model as an extension of the multi-dimensional Libor market model. In many cases, it is useful to work with approximating low-dimensional Markov models that allow for option pricing in finite difference grids. We shall here describe such a model which can be made approximately consistent with the cap and swaption expressions derived earlier.

As a starting point, let $f(t, T) = -\partial \ln P(t, T) / \partial T$ be the usual continuously compounded forward rate at time t for deposit over the interval $[T, T + dT]$. As shown by Heath, Jarrow, and Morton (HJM) (1992) any arbitrage free model with a single Brownian motion driving the yield curve level can be written as

$$df(t, T) = \mathbf{s}(t, T) \left[\int_t^T \mathbf{s}(t, s) ds dt + dW(t) \right]$$

where W is a scalar Brownian motion under the risk-neutral measure and $\{\mathbf{s}(t, T)\}_{T \geq t}$ is a collection of general stochastic processes adapted to the Brownian motion. The model is fully specified by the initial forward curve and a given volatility structure $\{\mathbf{s}(t, T)\}_{T \geq t}$. The

general one-factor HJM model requires as full continuum of all forward rates as state variables, with tree or lattice approximations of the model generally non-recombining and largely impractical.

Numerically tractable versions of the HJM model were treated in a range of papers in the early 90s, see for example Babbs (1993), Cheyette (1992), Jamshidian (1991), and Ritchken and Sankarasubrahmaniam (1992). In particular, it was noticed that when the forward rate volatility is of the separable form

$$\mathbf{s}(t, T) = g(T)h(t)$$

where g is a deterministic function and h is some adapted stochastic process, a finite state variable Markov representation of the yield curve is possible. Specifically, defining $\mathbf{n}(t) = -g'(t) / g(t)$ and $\mathbf{h}(t) = \mathbf{s}(t, t) = g(t)h(t)$ we get the bond reconstitution formula

$$P(t, T) = \frac{P(0, T)}{P(0, t)} e^{-G(t, T)x(t) - \frac{1}{2}G(t, T)^2 y(t)}, \quad G(t, T) = \int_t^T e^{-\int_t^u \mathbf{n}(s) ds} du \quad (6)$$

where, in the risk-neutral measure,

$$dx(t) = (-\mathbf{n}(t)x(t) + y(t)) dt + \mathbf{h}(t)dW(t), \quad x(0) = 0, \quad (7a)$$

$$dy(t) = (\mathbf{h}(t)^2 - 2\mathbf{n}(t)y(t)) dt, \quad y(0) = 0. \quad (7b)$$

As long as $\mathbf{h}(t) = \mathbf{h}(t, x(t), y(t))$, (6) and (7a-b) show that the resulting model is Markovian in only two state variables, x and y . The first state variable x can be interpreted as a yield curve factor, with the instantaneous short rate given by

$$r(t) \equiv f(t, t) = f(0, t) + x(t).$$

The second state variable y has the interpretation of a convexity correction term ensuring that the model is arbitrage-free. It is worth noting that the model only collapses to a single state variable model in the case when \mathbf{h} is deterministic. In this case the model corresponds to the general Gaussian model as presented in, for example, Jamshidian (1991).

In traditional implementations of the model above it is customary to let the volatility of the yield factor be a function of the short rate, for example

$$\mathbf{h}(t) \propto r(t)^p.$$

In this specification, however, the volatility skew induced by the coefficient p will tend to flatten for longer underlying tenors. To avoid this we here follow Andreasen (2000) and let \mathbf{h} be a function of a longer tenor forward swap rate, i.e. $\mathbf{h}(t) = \mathbf{h}(t, Y(t))$ where Y is some forward swap rate, the identity of which can change over time. Which swap rate is chosen at

any point in time will depend on the instrument that we wish to price. For example when pricing a standard Bermudan swaption² with final maturity T_e we let

$$Y(t) = R_{k,e}(t), \quad t \in [T_{k-1}, T_k[,$$

whereas when pricing a Bermudan callable cap we could set

$$Y(t) = R_{k,k+1}(t), \quad t \in [T_{k-1}, T_k[.$$

In any case, through (6), we would obviously have $\mathbf{h}(t) = \mathbf{h}(t, x(t), y(t))$. As a general rule, we notice that calibration and specification of low-dimensional Markovian models typically needs to be instrument-specific, unlike the Libor market models which are “large” enough to make possible calibration to the swaption and cap markets as a whole.

To incorporate stochastic volatility to our Markov model, we add to (7a-b) the third state variable V as given by (1b). The actual specification of our model is then

$$\mathbf{h}(t) = \sqrt{V(t)} \mathbf{I}(t) \mathbf{j}(Y(t)),$$

where \mathbf{I} is a deterministic function. As before, we will assume that $dZ \cdot dW = 0$.

For calibration of the model above to the swaption market closed-form or simple approximations for swaption prices are convenient. We note that under our assumptions the forward swap rate $R_{s,e}(t)$ evolves according to

$$dR_{s,e}(t) = \frac{\partial R_{s,e}(t)}{\partial x} \mathbf{h}(t) dW_{s,e}(t)$$

where $W_{s,e}$ is a scalar Brownian motion, and

$$\frac{\partial R_{s,e}(t)}{\partial x} = \frac{P(t, T_e)G(t, T_e) - P(t, T_s)G(t, T_s)}{A_{s,e}(t)} + R_{s,e}(t) \frac{\sum_{i=s+1}^e d_{i-1} P(t, T_i) G(t, T_i)}{A_{s,e}(t)} .$$

To price swaptions we approximate this derivative as being approximately deterministic³, and arrive at

$$dR_{s,e}(t) \approx \sqrt{V(t)} \mathbf{I}^*(t) \mathbf{j}(R_{s,e}(t)) dW_{s,e}(t), \quad (8)$$

$$\mathbf{I}^*(t) = \mathbf{I}(t) \left(\frac{\partial R_{s,e}(t)}{\partial x} \frac{\mathbf{j}(Y(t))}{\mathbf{j}(R_{s,e}(0))} \right)_{x(t)=y(t)=0} .$$

² One might question whether it is sufficient to use only a single driving Brownian motion for the yield curve when pricing Bermudan swaptions. Andersen and Andreasen (2001) conclude that the answer is generally yes, provided that calibration of the mean reversion parameter $\boldsymbol{\mu}$ is done carefully.

³ To understand why this is reasonable, we notice that if $R_{s,e}$ were the continuously compounded yield on a zero-coupon bond, it would be exactly linear in x .

We find that the approximation above provides sufficient accuracy for most applications, even for long-dated options out to, say, 30 years. Importantly, (8) allows us to use the results in Sections 3 and 4 to efficiently price swaptions and caps. Due to the structure of the model, once the mean-reversion of rate \mathbf{n} has been set⁴, the model volatilities $\{\mathbf{I}(t)\}$ can be bootstrap-calibrated to match a strip of swaption prices. On a standard PC, bootstrap calibration to, say, the swaption strip $\{1 \times 29, 2 \times 28, \dots, 29 \times 1\}$ can be done in less than 5 seconds when using transform inversion and in a few tenths of second using expansions (for comparison, the calibration can be done in less than one tenth of a second when volatilities are deterministic). Figure 2 demonstrates the quality of the fit for the case of $\mathbf{j}(x) = mx + (1 - m)L$, $m = 0.2$.

6. Numerical Methods.

Numerical implementation of the general Libor market model (1a-b) presented in Section must virtually always be done through Monte Carlo simulation. Due to the zero correlation between the forward rate processes and the stochastic volatility process, it is possible to split the Monte Carlo generation of interest rate paths into two pieces: 1) draw a path of the variance process V through time; 2) draw a path of forward rates assuming that $\mathbf{I}_k(t)\sqrt{V(t)}$ is deterministic. Schemes for step 2) are well-known (see e.g. Andersen and Andreasen (2000)), so we here focus on simulating (1b) on some discrete time-grid $\{t_i\}_{i=0,1,\dots}$.

A direct Euler or log-Euler discretization of (1b) is prone to instability unless either the time step or the mean-reversion parameter \mathbf{k} are small. Noticing the exact result

$$E(V(t_{i+1}) | V(t_i)) = \mathbf{q} + (V(t_i) - \mathbf{q})e^{-\mathbf{k}(t_{i+1}-t_i)}$$

it is generally better to resort to a moment matching scheme where, for instance, we can approximate $V(t_{i+1}) | V(t_i)$ as a log-normal variable:

$$\hat{V}(t_{i+1}) = \left(\mathbf{q} + (\hat{V}(t_i) - \mathbf{q})e^{-\mathbf{k}(t_{i+1}-t_i)} \right) e^{\frac{1}{2}\Gamma(t_i)^2 + \Gamma(t_i)\tilde{z}_i},$$

$$\Gamma(t)^2 = \ln \left(1 + \frac{\frac{1}{2}\mathbf{e}^2 \mathbf{y} (\hat{V}(t_i))^2 \mathbf{k}^{-1} (1 - e^{-2\mathbf{k}(t_{i+1}-t_i)})}{(\mathbf{q} + (\hat{V}(t_i) - \mathbf{q})e^{-\mathbf{k}(t_{i+1}-t_i)})^2} \right),$$

where the \tilde{z}_i 's are a sequence of i.i.d. standard Gaussian draws.

Monte Carlo schemes for the one-factor Markovian model presented in Section 5 are similar to those of the full LM model. Importantly, however, the limited number of state variables (a total of three) allows us to write down a low-dimensional PDE for option prices

⁴ To prevent the model from being too non-stationary, it is often best to let the rate mean reversion be constant. For instance, we could best-fit \mathbf{n} to reproduce the auto-correlation of the swap rates as computed approximately in a globally fitted LM yield curve model; see Andreasen (2000) for details.

which can feasibly be solved on a computer. Specifically, any contingent claim Υ satisfies, subject to boundary conditions, the equation

$$\left(\frac{\partial}{\partial t} + D_x + D_y + D_V \right) \Upsilon = 0,$$

where

$$\begin{aligned} D_x &= -\frac{1}{3}r + (-\mathbf{n}x + y) \frac{\partial}{\partial x} + \frac{1}{2}\mathbf{h}^2 \frac{\partial^2}{\partial x^2}, \\ D_y &= -\frac{1}{3}r + (\mathbf{h}^2 - 2\mathbf{n}y) \frac{\partial}{\partial y}, \\ D_V &= -\frac{1}{3}r + \mathbf{k}(\mathbf{q} - V) \frac{\partial}{\partial V} + \frac{1}{2}\mathbf{e}^2 \mathbf{y}(V)^2 \frac{\partial^2}{\partial V^2}, \end{aligned}$$

with $r = f(0, t) + x$.

The three-dimensional PDE above can be handled numerically using a number of available solvers. One good method which takes advantage of the fact that the PDE contains no mixed derivatives is the so-called Alternating Directions Implicit (ADI) finite difference method; see for instance the 3-dimensional Douglas scheme described Mitchell and Griffiths (1980). This scheme is uniformly stable, has the same order of computational effort as a multinomial tree, yet allows for complete freedom in grid design and has second order accuracy in the time-domain (as opposed to trees that only have first order time accuracy). Non-zero cross terms, i.e. correlation between the rates and the volatility, can be incorporated without loss of accuracy or stability but at a cost of making the scheme computationally more expensive⁵. In our implementation of the ADI scheme for the PDE above we use a standard three-point discretisation for the x - and V -dimensions, but a five-point discretisation for the y -dimension. This enables us to use a moderate number of grid points in the y dimension, around 10 or so, without sacrificing much accuracy. On a standard PC a 30-year Bermudan swaption can be priced in about 10 seconds in the stochastic volatility model. Switching off the stochastic volatility reduces the calculation time to about 0.5 seconds for the same contract.

7. Some empirical observations.

Estimation of the process (1b) is most easily done from observations of movements of implied volatilities (or their squares, the implied variances) of caps and swaptions. In particular, we notice that estimation needs to be done under a pricing measure, not the “real” historical measure, which makes traditional time-series estimation of historical volatility largely useless for our purposes.

The process (1b) implies that volatility of implied variance should decay with increasing option maturity, eventually reaching an asymptotic level. The asymptotic level is a

⁵ In certain special cases it is possible to change variables locally to eliminate the cross-terms, resulting in only a fairly moderate increase in computation time. More generally, we would need to modify the numerical scheme to explicitly incorporate correlation terms. Typical schemes for this (e.g. predictor-corrector schemes) would increase computation time by a factor of around two.

function of the skew specification \mathbf{j} ; the speed with which the volatility of implied variance approaches the asymptote is a function of the mean reversion parameter \mathbf{k} . Figure 3 demonstrates that this behavior is empirically observable for various swap tenors in USD. It is, however, obvious from the figure that the decay towards the asymptote is faster for short-tenor rates than for long-tenor ones. To make our model consistent with this behavior we find that we generally need mean-reversions \mathbf{k} of around 0.3-0.4 for medium-to-long-tenor rates in USD⁶, whereas short-dated rates (say, with tenors less than a year) are better fit with mean reversions of 1 or higher. It is also obvious from the figure that the volatility of variance parameter \mathbf{e} is higher for short-tenor rates than for long-tenor rates; in the US we find that rates with tenors less than 1-year typically require $\mathbf{e} \approx 150\%$, whereas longer-tenor swap rates are best described with $\mathbf{e} \approx 100\%$ or less. The tenor-dependence of parameters can likely be explained as a diversification effect, as the variance processes of different forwards are, of course, less than perfectly correlated (which is what we effectively assume in our model with a single stochastic volatility factor). Still, we find empirically that movements of implied volatilities of most swap forward rates are quite high (typically above 80%, at least when the rate tenors exceed 3-5 years) and as such consider our setup a reasonable representation of reality. It is in principle not difficult to extend our LM model framework to multi-factor variance processes, but, as discussed earlier, the difficulties of populating the parameters of the model makes this of somewhat limited practical appeal.

Turning now to the question of how well stochastic volatility models can reproduce the observed volatility skews and smiles, we generally find that the model does a good job in both cap and swaptions markets. In Figure 4 we show typical fits for EUR and USD swaptions data. We typically find that parameters estimated from the historical decay-properties of volatility of implied variance are close to those required to match observed volatility smiles. In other words, out-of-the-money swaptions are priced approximately at the cost of vega-hedging with at-the-money swaptions.

8. Pricing Simple Exotics.

When applying stochastic volatility models to simple instruments, one can often get inspiration from the structure of the one-factor Markov model in Section 5 to come up with reasonable approximations. This, for instance, is useful when pricing European options on amortizing swaps, or options and swaps where the floating rate is a long-dated swap rate (so-called CMS rate). To illustrate this, consider the time 0 value $M(0)$ of a CMS coupon paying a time⁷ T_s the swap rate $R_{s,e}(T_s)$:

$$M(0) = P(0, T_s) E^s \left(R_{s,e}(T_s) \right) = P(0, T_s) E^{A_{s,e}} \left(R_{s,e}(T_s) \frac{A_{s,e}(0) / P(0, T_s)}{A_{s,e}(T_s)} \right) \quad (9)$$

where $E^s(\cdot)$ as before denotes expectation under the martingale measure with the maturity T_s zero-coupon bond as numeraire.

⁶ In EUR, we generally find that mean reversions are slightly lower than in USD. See Figure 4.

⁷ We here ignore the fact that actual payments are typically made in arrears, that is at time T_{s+1} . The adjustment for this payment delay can be performed using the same technique as shown here.

To evaluate (9) we need to be able to say something about the distribution of $A_{s,e}^{-1}(T_s)$ under the swap annuity measure. The idea is now to come up with an inspired approximation for the annuity $A_{s,e}(T_s)$ as function of the terminal swap rate $R_{s,e}(T_s)$ and compute the expectation in (9) using earlier results. Dropping the subscripts on A and R , we first note that

$$E^A(y(T_s)) \equiv \bar{y}(T_s) \approx \text{var}^A(x(T_s)) \approx \text{var}^A(R(T_s)) \left(\frac{\partial R(T_s)}{\partial x} \Big|_{x(T_s)=y(T_s)=0} \right)^{-2}$$

where $\text{var}^A(R(T_s))$ is the variance in the annuity measure which can be implied from the swaption model of the first section. The swap rate derivative in the expression above can be obtained from a chosen level of rate mean-reversion \mathbf{n} .

We can now use the reconstruction formula (6) to write

$$\begin{aligned} R(T_s) &= R(T_s, x(T_s), y(T_s)) \approx R(T_s, x(T_s), \bar{y}(T_s)); \\ A(T_s) &= A(T_s, x(T_s), y(T_s)) \approx A(T_s, x(T_s), \bar{y}(T_s)). \end{aligned}$$

As the first of these functional relationships is one-to-one, we get an expression $A(T_s) \approx l(T_s R(T_s))$, for some explicitly computable function l . To improve the approximation, we actually use

$$A(T_s) = c \cdot l(T_s R(T_s)) \tag{10}$$

where the constant c is chosen so that the Radon-Nikodym satisfies the natural condition

$$E^A \left(\frac{A(0)/P(0, T_s)}{c \cdot l(T_s R(T_s))} \right) = 1$$

Equation (10) can now be inserted in (9) which again can be integrated on the density of R to value the CMS coupon. The density of R can, as always, be obtained by differentiation of swaption prices:

$$\text{Prob}^{A_{s,e}}(R_{s,e}(T_s) \in [K, K + dK]) / dK = \partial^2 S_{s,e}(0; K) / \partial K^2 \cdot A_{s,e}(0)^{-1}$$

where $S_{s,e}(0; K)$ is the time 0 price of a European payer swaption struck at K . Expressions for these prices have been derived in Sections 3 and 4. In Figure 5, we examine the performance of the above expression for different levels of mean-reversion \mathbf{n} . Notice that the convexity adjustments increase with mean reversion, a consequence of increased volatility of the short end of the curve (and thereby of the annuity A which has relatively short duration).

9. Pricing Exotics and Callable Instruments.

The techniques discussed in Section 8 are only useful for simple, European instruments. For more complicated instruments, such as Bermudan swaptions and other callable or path-dependent structures, the numerical methods discussed in Section 7 must be applied. For strongly path-dependent options, Monte Carlo simulation of the full Libor market model is the method of choice whereas Bermudan swaptions are probably best handled in a the low-dimensional Markov model coupled with an ADI finite difference solver. We do point out, however, that methods to price Bermudan/American options inside a Monte Carlo simulation of the full Libor market model exist (see e.g. Andersen and Broadie 2001 for a review) and will likely become more prevalent in the future as available computing power continues to increase.

In Table 1 we use the ADI method to compare prices for 30 year Bermudan swaptions in the stochastic volatility model with those of the deterministic volatility model. Provided that the deterministic volatility model is calibrated to the European swaptions with strike equal to that of the underlying swap, the differences between Bermudan prices in the stochastic and deterministic volatility model are, perhaps suprisingly, not particularly significant.

References

- Andersen, L. and M. Broadie (2001): "A practical primal-dual simulation algorithm for high-dimensional American options," Working Paper, Gen Re Securities and Columbia University.
- Andersen, L. and J. Andreasen (2000): "Volatility Skews and Extension of the Libor Market Model," *Applied Mathematical Finance*, 7, 1-32.
- Andersen, L. and Andreasen, J. (2001). "Factor-Dependence of Bermudan Swaptions: Fact or Fiction?" *Journal of Financial Economics*, 62, 3-37.
- Andersen, L. and R. Brotherton-Ratcliffe (2001): "Extended Libor Market Models with Stochastic Volatility," Working paper, Gen Re Securities. www.ssrn.com.
- Andreasen, J. (2000): "Turbo-charging the Cheyette Model," Working paper, Gen Re Securities.
- Babbs, S. (1993): "Generalised Vasicek' Models of the Term Structure." *Applied Stochastic Models and Data Analysis*, Vol 1, 49-62.
- Cheyette, O. (1992): "Markov Representation of the Heath-Jarrow-Morton Model." Working paper, BARRA.
- Glasserman, P. and S. Kou (1999). "The term structure of simple forward rates with jump risk," Working Paper, Columbia University

Heath, D., Jarrow, R. and A. Morton (1992): "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation." *Econometrica* **60**, 77-106.

Heston, S. (1993): "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options." *Review of Financial Studies* **6**, 327-344.

Jamshidian, F. (1991): "Bond and Option Evaluation in the Gaussian Interest Rate Model." *Research in Finance* **9**, 131-710.

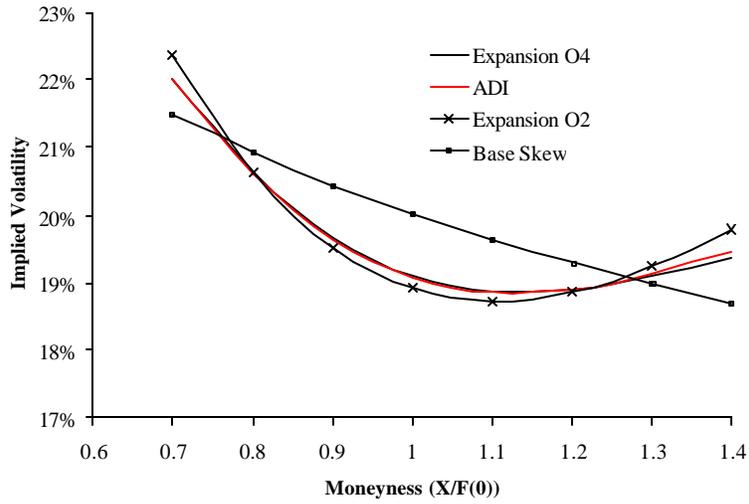
Lewis, A. (2000): *Option Valuation under Stochastic Volatility*. Finance Press.

Mitchell, A. and D. Griffiths (1980): *The Finite Difference Method in Partial Differential Equations*. John Wiley, New York.

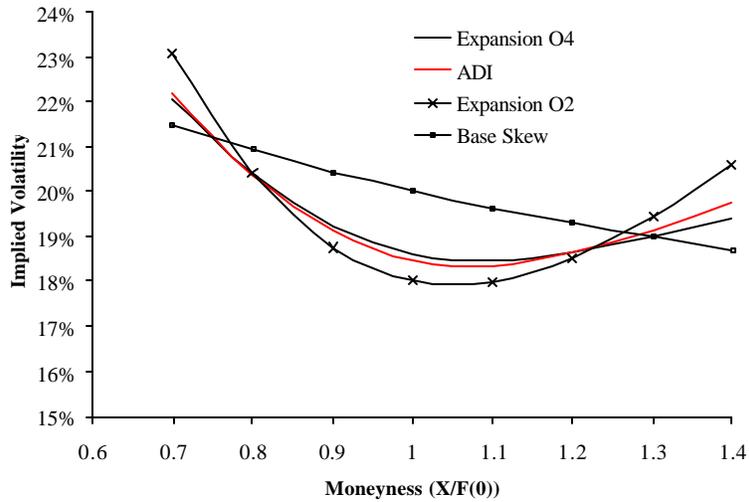
Rebonato, R. (2001), "The stochastic volatility LIBOR market model," *RISK*, October, 105-110.

Ritchken, P. and Sankarasubramaniam (1993): "On Finite State Markovian Representations of the Term Structure." Working paper, Department of Finance, University of Southern California.

Figure 1: Implied 1.5-year Caplet Volatility Smile



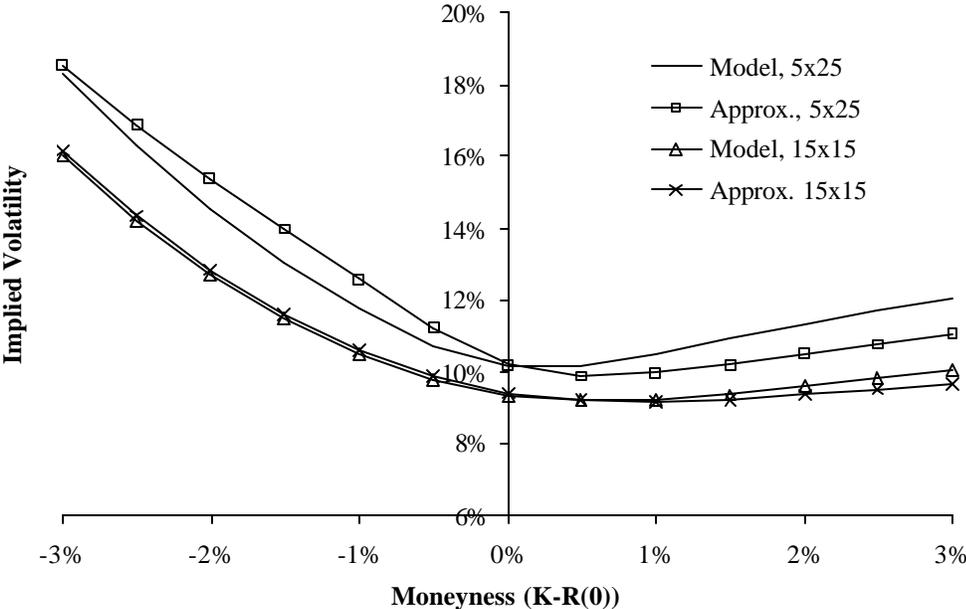
Panel A: $e = 150\%$, $k = 1$



Panel B: $e = 200\%$, $k = 1$

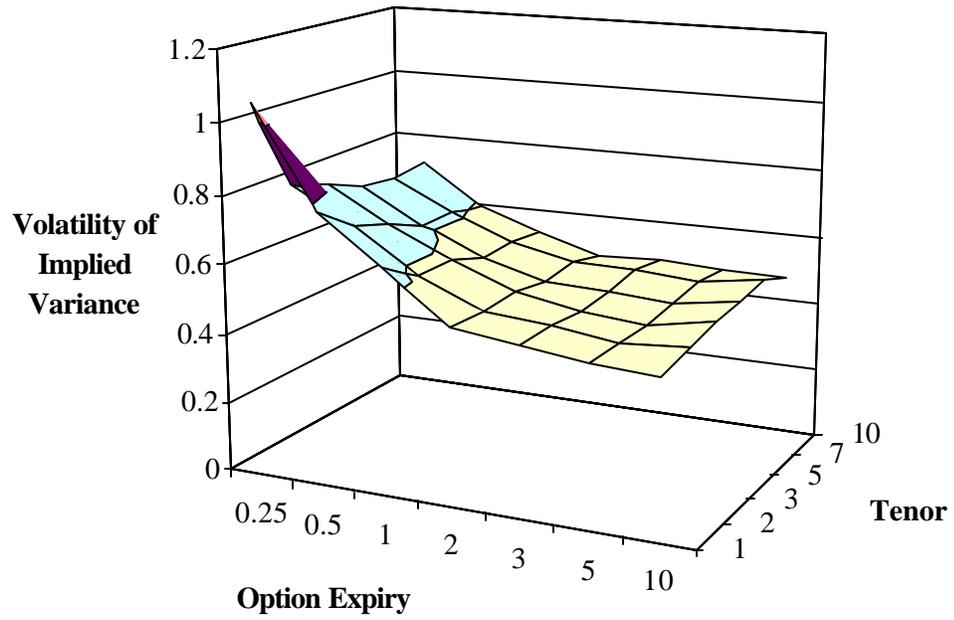
Notes: Panels A and B show the implied (Black) volatility smile for a 1.5-year caplet with $k = 1$ and $e = 150\%$ and 200% , respectively. The initial forward was 6% , $V(0) = q = 1$, and the model skew functions were $j(x) = x^{0.6}$ and $y(x) = x^{3/4}$. The volatility (σ) was set to 6.49% (around 20% in log-normal terms). In the computation of expansion O4 results, the cut-off parameter Λ in Eq. (21) was set to 1 . The "ADI" numbers reported in the graph were computed in a two-dimensional finite difference grid; the "base skew" numbers correspond to $e = 0$ (no stochastic volatility component).

Figure 2: Swaption Volatility Smile. Yield Curve Model versus Approximation



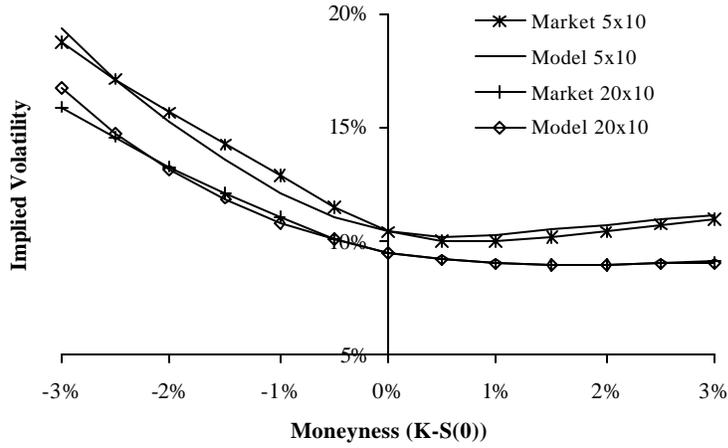
Notes: Sample of implied Black swaption volatility smiles of yield curve model (“Model”) in Section 5, vs. the approximation eq. (8). The yield curve model is calibrated to ATM European swaption prices in EUR (February, 2002) for the strip $1 \times 29, 2 \times 28, \dots, 29 \times 1$. The graph use the simple displaced diffusion skew function $j(R(t)) = mR(t) + (1 - m)R(0)$, $m = 0.2$, and $y(x) = \sqrt{x}$. Other process parameters were $e = 1, k = 0.1, V(0) = q = 1$. In the full yield curve model prices are found by finite-difference solution.

Figure 3: Volatility of Implied Variance of Various Rates in US

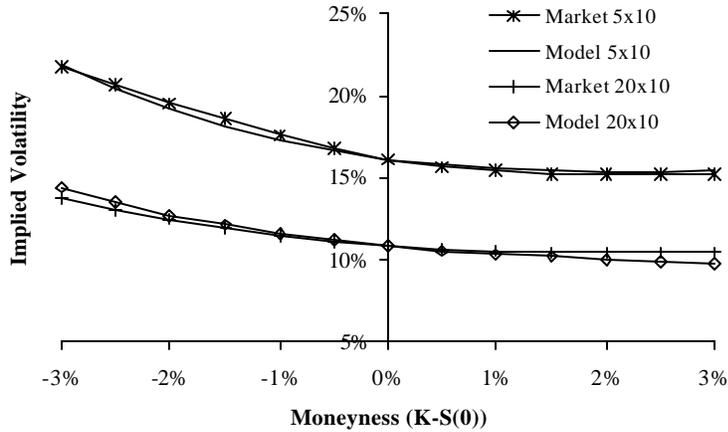


Notes: The graphs shows the volatility of implied variance (= the square of implied volatility) implied for US swaptions on 1-, 2-, 3-, 5-, 7-, and 10-year rates, as estimated from two years of weekly data from Bloomberg.

Figure 4: Volatility Smile of Stochastic Volatility Model vs. Market



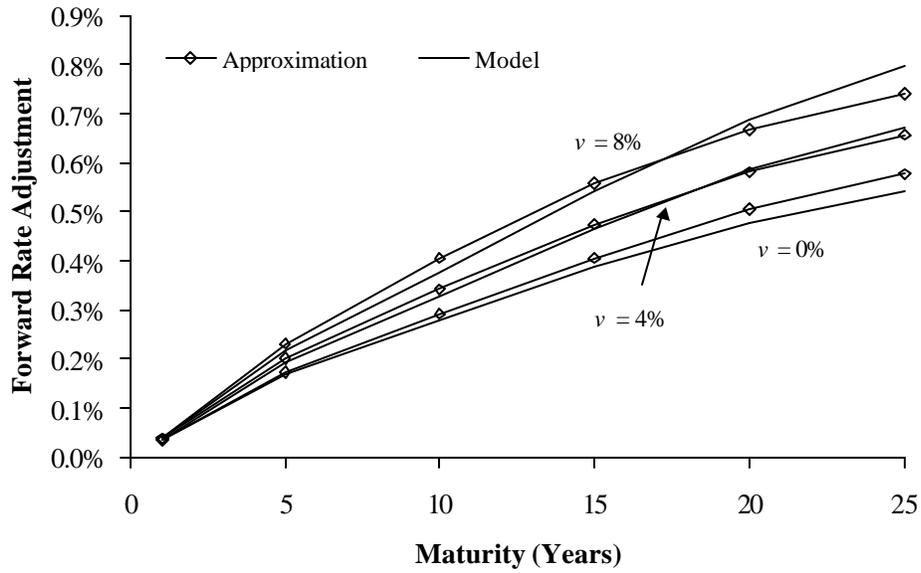
Panel A: EUR data, February 2002



Panel B: USD data, February 2002

Notes: the graphs show swaption smile as implied by a stochastic volatility model (2) and as observed in the market. The graph use the simple displaced diffusion skew function $j(R(t)) = mR(t) + (1-m)R(0)$ and $y(x) = \sqrt{x}$. Parameters of the stochastic volatility process in USD were $m = 0.3$, $e = 1$, $k = 0.3$; in EUR the parameters were $m = 0$, $e = 1$, $k = 0.1$. In both currencies we set $V(0) = q = 1$.

Figure 5: CMS Adjustment as Function of Expiry in EUR



Notes: CMS forward rate adjustment as function of maturity and rate mean reversion ν . The rate tenor is 20 years and the market is EUR, February 2002. The “Model” numbers were computed in a finite difference grid; the “Approximation” numbers were computed as outlined in Section 8. Model parameters were as in Figure 4, Panel A.

Table 1: Bermudan Prices in EUR under Deterministic and Stochastic Volatility

strike	30nc1 par receiver			30nc1 par payer			30nc1 accr receiver		
	sv	dv	vega diff	sv	dv	vega diff	sv	dv	vega diff
3.0%	1.55%	1.46%	0.40%	36.01%	36.29%	-1.31%	2.56%	2.38%	0.50%
3.5%	2.04%	2.07%	-0.08%	29.65%	30.12%	-1.50%	3.63%	3.62%	0.02%
4.0%	2.72%	2.83%	-0.30%	23.68%	24.35%	-1.45%	5.18%	5.32%	-0.21%
4.5%	3.67%	3.82%	-0.31%	18.27%	18.99%	-1.24%	7.45%	7.67%	-0.23%
5.0%	5.03%	5.14%	-0.18%	13.67%	14.17%	-0.69%	10.82%	10.98%	-0.12%
5.5%	7.09%	7.06%	0.04%	10.05%	10.08%	-0.04%	15.82%	15.86%	-0.03%
6.0%	10.21%	10.12%	0.12%	7.39%	7.15%	0.25%	23.18%	23.14%	0.03%
6.5%	14.74%	14.73%	0.01%	5.51%	5.29%	0.23%	33.86%	34.31%	-0.26%
7.0%	20.37%	20.48%	-0.28%	4.18%	3.91%	0.30%	48.56%	49.11%	-0.39%
7.5%	26.65%	26.75%	-0.40%	3.25%	3.01%	0.30%	67.13%	67.61%	-0.46%
8.0%	33.25%	33.32%	-0.49%	2.56%	2.36%	0.28%	89.38%	89.85%	-0.61%

Notes: Prices of EUR Bermudan swaptions (February 2002) under stochastic (“sv”) and deterministic (“dv”) volatility. “30nc1” refers to an annual Bermudan swaption of 30 year maturity with the first exercise date 1 year from now. “par” means that the underlying swap of the standard par type, whereas “accr” means that the underlying swap is of the accreting (or zero-coupon) type. Vega differences (“vega diff”) refer to the difference in sv and dv option prices expressed in Black implied volatility terms. The model used is the one-factor Markov model in Section 5 with skew specification of $\mathbf{j}(R(t)) = mR(t) + (1-m)R(0)$. For the deterministic volatility model we use $m=0$ for the strikes less than or equal to 5.5%, and $m=1$ for the remaining strikes. In all cases the deterministic volatility model is calibrated to the European swaptions $1 \times 29, 2 \times 28, \dots, 29 \times 1$ with the same strike as the Bermudan; the Europeans are valued in a volatility model of type eq. (2). For the stochastic volatility model we use the 1-factor model in Section 5 with $m=0$ and $\mathbf{y}(x) = \sqrt{x}$, and calibrate the model to the at-the-money swaptions $1 \times 29, 2 \times 28, \dots, 29 \times 1$. The stochastic volatility parameters are as in Figure 4, Panel A. All reported prices are found by finite difference solution.