Extended Libor Market Models with Stochastic Volatility[1]

Leif Andersen
Rupert Brotherton-Ratcliffe
Gen Re Securities

First Version: May 2001
This Version: December 2001

Abstract

This paper introduces stochastic volatility to the Libor market model of interest rate dynamics. As in Andersen and Andreasen (2000a) we allow for non-parametric volatility structures with freely specifiable level dependence (such as, but not limited to, the CEV formulation), but now also include a multiplicative perturbation of the forward volatility surface by a general mean-reverting stochastic volatility process. The resulting model dynamics allow for modeling of non-monotonic volatility smiles while explicitly allowing for control of the stationarity properties of the resulting model dynamics. Using asymptotic expansion techniques, we provide closed-form pricing formulas for caps and swaptions that are robust, accurate, and well-suited for both model calibration and general mark-to-market of plain-vanilla instruments. Monte Carlo schemes for the proposed model are proposed and examined.

Keywords: volatility smiles, stochastic volatility, Libor market model, asymptotic expansions, ADI finite differences, Monte Carlo simulation

1. Introduction.

The Libor market (LM) model of Jamshidian (1997), Brace *et al* (1997), and Miltersen *et al* (1997) has established itself as an important model for pricing and hedging of fixed income derivatives. The model is flexible, supports multiple factors and rich volatility structures, and is tractable enough is to allow fast calibration to market-quoted caps and swaptions. In its original log-normal form, however, the LM model does not incorporate the observable phenomenon that implied log-normal volatilities of quoted interest rate caps and swaptions are strike- and coupon-dependent, respectively. This is a fairly significant drawback of the basic model, as virtually all fixed income markets exhibit significant downward-sloping volatility "skews" (i.e. cap volatilities decrease with strike) with some markets even displaying non-monotonic "smiles" where both in-the-money and out-of-the-money options trade at higher implied volatilities than at-the-money options.

[1] The authors wish to express their gratitude to Jesper Andreasen and Alan Lewis for discussions and suggestions.

To incorporate volatility skews and smiles into the LM model, researchers have suggested a variety of extensions of the basic framework. Andersen and Andreasen (2000a) modify the assumption of log-normal diffusion to allow for local forward rate volatilities that have arbitrary dependence on the forward rate itself. Such effects are empirically observable, see for instance the US data in Figure 0 below. Even stronger level dependence can be observed in Japanese markets.

Figure 1 Here

The introduction of level dependency of forward rate volatilities in LM models essentially corresponds to the deterministic volatility function (DVF) approach by Dupire (1994) for equity options, and works well for markets where implied volatilities are monotonic functions of the option strike. For the case of non-monotonic volatility "smiles", the approach is, however, prone to non-stationarities in the evolution of implied volatility and is generally less useful (see for instance Andersen and Andreasen (2000b) for a discussion of this phenomenon in an equity setting).

In another line of research, Glasserman and Kou (1999) and Jamshidian (1999) instead extend the Libor market model to incorporate jumps in forward rates curves. This work is related to that of Merton (1976) and in certain special cases (e.g. log-normally distributed jumps) has enough tractability to allow for closed-form solutions for European interest rate options. While the jump approach can generate stationary, non-monotonic volatility smiles, it involves a variety of technical complications, and numerical execution of the resulting model is non-trivial. Moreover, to generate asymmetric smiles and skews, the jump component of the forward rate dynamics typically needs to be of substantial magnitude, typically with a downward bias. While such dynamics are probably reasonable for equity prices (see Andersen and Andreasen (2000b)) they might be less natural for the term structure of interest rate forwards.

The approach to volatility smile modeling discussed in this paper involves overlaying the level-dependent volatility function approach in Andersen and Andreasen (2000a) with an independent stochastic volatility process. Empirical evidence strongly supports the qualitative basis of such a model, as volatilities in interest rate options markets clearly possess a random component beyond that implied by the DVF approach. Application of stochastic volatility models to equity option pricing[2] is common and has received much scrutiny throughout the last decade or so (e.g. Hull and White (1987), Heston (1993), and the excellent monograph by Lewis (2000)), but comparatively little work has been done on applications to problems in fixed income. Of direct relevance to our work is primarily the recent paper by Joshi and Rebonato (2001) who investigate a

[2] Note that virtually all of this literature assumes that the stochastic volatility process is imposed on a regular, skew-free geometric Brownian motion process.

displaced[3] log-normal LM model with a particular four-parameter parameterization of the instantaneous forward rate volatility term structure. To introduce stochastic volatility, each of the four parameters is assumed to follow an Ornstein-Uhlenbeck process (or a transformation thereof), with all processes being independent of each other and of the forward curve. While the assumption of independence allows for some relief of computational burden, the model is not analytically tractable and caps and swaptions must be priced by Monte Carlo simulation.

The approach taken in this paper differs from that of Joshi and Rebonato (2001) in several important ways. We allow for a richer class of forward rate diffusion processes, making it possible for the model builder to establish virtually any kind of skew/smile "bedrock" on which to superimpose the stochastic volatility process. This has several consequences, among others the ability to potentially sustain smiles for long maturities, as well as the ability to explicitly control how much of a non-monotonic volatility smile should be stationary (in a sense defined shortly) and how much should not. Also, we allow for complete freedom in the specification of the instantaneous volatility term structure, accommodating both parametric and non-parametric calibration approaches (for a discussion of the latter, see for instance Sidenius (2000) and Andersen and Andreasen (2001)). While we allow for several extensions, our basic approach to stochastic volatility modeling is parsimonious with a single, mean-reverting stochastic factor perturbing the volatility term structure. Most importantly, we provide accurate closed-form asymptotic expansions for European cap and swaption prices, ensuring that the proposed model allows for fast calibration and is suitable as a production model to mark-to-market cap and swaption positions. This feature is, of course, of importance in preventing model arbitrages between plain-vanilla and exotic options. We note that applications of asymptotic techniques to derivatives pricing problems involving stochastic volatility are relatively well-established (see e.g. Hull and White (1987), Hagan (2000), and Lewis (2000)), but previous work has not been general enough for our purposes.

The rest of the paper is organized as follows. Section 2 provides notation and outlines the basic model assumptions. Section 3 uses a variety of asymptotic techniques to derive an expansion for the prices of caps and floors. In Section 4, the results of Section 3 are applied to European swaption pricing, for which an approximation is found. Section 5 tests the cap and swaption pricing formulas against numerical results from Monte Carlo simulations (an algorithm for which is outlined in Appendix A) and a 2-factor alternating directions implicit (ADI) finite difference grid. See In Section 6 we briefly examine implied volatility data from the US market and demonstrate that the proposed framework is capable of capturing important characteristics of the implied

[3] A displaced log-normal martingale x satisfies $dx(t) = ...dt + [a + bx(t)]dW(t)$ where a and b are positive constants and where W is a Brownian motion. We notice that this process allows for negative values of x; the domain of x is $[-a/b, \infty)$.

volatility dynamics. We also propose a number of extensions of the basic model framework. Section 7 contains the conclusions of our paper.

2. Model setup, assumptions, and notation.

Consider an increasing maturity structure $0 = T_0 < T_1 < ... < T_{K+1}$ and define a right-continuous mapping function $n(t)$ by $T_{n(t)-1} < t \le T_{n(t)}$. With $P(t,T)$ denoting the time t price of a zero-coupon bond maturing at time T, we define discrete forward rates (*Libor forward rates*) on the maturity structure as follows (where $t \le T_k$ and $k \le K$):

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

We wish to state the dynamics of the Libor forward rates in the spot measure **Q**, the equivalent martingale measure induced by a discrete-time Libor money market account B generated by "rolling over" an initial investment of \$1 at each date in the maturity structure. Specifically,

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

Under the extended LM model in Andersen and Andreasen (2000a), the no-arbitrage dynamics of forward rates in **Q** are governed by the following set of stochastic differential equations, $k = n(t),\ldots,K$:

$$dF_k(t) = \gamma\left(F_k(t)\right)\lambda_k(t)^{\mathrm{T}}\left[\mu_k(t)dt + dW(t)\right], \quad \mu_k(t) = \sum_{j=n(t)}^{k} \frac{\delta_j \gamma\left(F_j(t)\right)\lambda_j(t)}{1 + \delta_j F_j(t)} \ . \qquad (2)$$

In (2), $\gamma : \mathbb{R}^+ \to \mathbb{R}^+$ is a one-dimensional function satisfying certain regularity conditions, $\lambda_k(t)$ is a bounded m-dimensional deterministic function, and $W(t)$ is a m-dimensional Brownian motion under **Q**. (2) defines a system of up to K Markov state variables. The setup in (2) allows for robust modeling of downward-sloping volatility skews, for instance by using a CEV model with $\gamma(x) = x^\alpha$, $0 < \alpha < 1$ for which a closed-form option pricing formula is available. For other specifications of γ, one can employ either finite-difference grids or, more conveniently, the asymptotic pricing formulas derived in Section 3.1.

In principle, the specification (2) can also be used to generate non-monotonic, *U*-shaped volatility "smiles" by letting $\gamma(x)/x$ be non-monotonic. It is clear, however, that such a model would often imply quite non-stationary behavior of the volatility smile as a function of forward rate levels. For instance, consider the "double" CEV model

$\boldsymbol{j}(x) = x^{a} + wx^{b}$, $0 < \boldsymbol{a} < 1$, $\boldsymbol{b} > 1$, with w calibrated to generate a volatility smile bottoming out at current forward rate levels. If forward rates were to move up considerably from calibration levels, the "super-lognormal" x^{b} term would dominate, resulting in an inverted volatility skew with caplet volatilities increasing in strike. This behavior does not correspond well to observed market behavior where the smile tends to "slide along" with the forward rate. Ideally, we would like to include in our model a mechanism that would complement our existing framework by allowing for such behavior.

Generation of stationary non-monotonic volatility smiles can be accomplished by allowing the underlying forward rates to jump Poisson-style, or by introducing stochastic volatility (or possibly a combination of both). As discussed earlier, we find the stochastic volatility approach to be the most appealing, and will focus on this throughout this paper. We will take the process (2), likely equipped with a $\boldsymbol{j}$ that generates either a flat or monotonically down-ward sloping volatility skew (or at least only a mildly non-monotonic smile), as our starting point, but allow the term on the Brownian motion to be scaled by a stochastic process. Specifically, we introduce a mean-reverting scalar[4] process

$$dV(t) = \boldsymbol{k}\left(\boldsymbol{q} - V(t)\right)dt + \boldsymbol{e}\boldsymbol{y}\left(V(t)\right)dZ(t) \tag{3}$$

where $\boldsymbol{k}, \boldsymbol{q}$, and $\boldsymbol{e}$ are positive constants, $Z(t)$ is a Brownian motion under $\mathbf{Q}$, and $\boldsymbol{y} : \mathbb{R}^{+} \to \mathbb{R}^{+}$ is a well-behaved function. We obviously must impose that (3) will not generate negative values of V, which requires $\boldsymbol{y}(0) = 0$. We will interpret the process in (3) as the (scaled) variance process for our forward rate diffusions. Specifically, the square-root of the process in (3) will be used as a stochastic, multiplicative shift of the diffusion term in (2). In total, our process assumption is:

Assumption 1:
The forward rate process under $\mathbf{Q}$ *is*

$$dF_{k}(t) = \boldsymbol{j}\left(F_{k}(t)\right)\sqrt{V(t)}\,\boldsymbol{l}_{k}(t)^{\mathrm{T}}\left[\sqrt{V(t)}\boldsymbol{m}_{k}(t)dt + dW(t)\right], \tag{4}$$

where $\boldsymbol{m}_{k}(t)$ *is given in (2) and where* $V(t)$ *follows (3).*

In applications, it is often natural to scale the process for V such that $\boldsymbol{q} = 1$. The quantity $1 - V(t)$ then represents the percentage at which current levels of V deviate from the long-term mean. Also, we note that the inverse mean reversion parameter $\boldsymbol{k}^{-1}$ has units of time and represents the half-life of volatility shock decay. So, if (in the

[4] Extensions to multi-dimensional variance processes are possible; see Section 6.

pricing measure $\mathbf{Q}$) we expect the effects of a volatility shock to be reduced by 50% within 6 months, we must set $k = 2$.

We shall also need the following:

Assumption 2:
The Brownian motion of the variance process (3) is uncorrelated to the Brownian motions of the forward curve, i.e.

$$\text{diag}\left(dZ(t)\right) \cdot dW(t) = \mathbf{0} \ .$$

Assumption 2 might, at first glance, appear somewhat prohibitive. However, empirical evidence from all major fixed income markets generally suggests that correlations between short-dated forward rates and their volatilities are indistinguishable[5] from 0; see e.g. the analysis in Chen and Scott (2001). In traditional models for stochastic volatility (such as the Heston (1993) model) built on top of geometric Brownian motion, non-zero correlations are often necessary to fit observed asymmetric volatility skews in the market. In our more general set-up, however, the base skew is established by the function $\jmath$ and non-zero correlations are not needed to generate asymmetric skews. Loosely speaking, the effects of non-zero correlation in a geometric Brownian model with stochastic volatility are here generated by two separate pieces: a zero-correlation piece (the V process) and a piece that is either perfectly dependent or perfectly anti-dependent on the forward process (the effect of the skew function $\jmath$).

The main technical implication of Assumption 2 is the ability to shift pricing numeraires without affecting the process (3). For example, consider substituting the numeraire $B(t)$ for a specific zero-coupon bond $P(t,T_{k+1})$, with $k \le K$ fixed. Let the martingale measure induced by this choice of numeraire be $\mathbf{Q}^{k+1}$. The Radon-Nikodym derivative for the shift from $\mathbf{Q}$ to $\mathbf{Q}^{k+1}$ is thus

$$\frac{d\mathbf{Q}^{k+1}}{d\mathbf{Q}}\Big|\Im_t = \frac{P(t,T_{k+1}) \,/\, P(0,T_{k+1})}{B(t)} \equiv \rho(t), \ t \le T_{k+1}.$$

Letting $< \cdot, \cdot >$ define quadratic co-variance, the dynamics in the $\mathbf{Q}^{k+1}$ measure are then:

$$dF_k(t) = \jmath\left(F_k(t)\right)\sqrt{V(t)}\,\lambda_k(t)^{\mathrm{T}}\,dW_{k+1}(t)\,, \qquad (5a)$$

[5] In our framework, movements of implied volatilities of forward rates are caused both by the random nature of V and the presence of the skew function $\jmath$. Over short time-horizons – as in the study by Chen and Scott (2001) – the first effect dominates. See Section 6 for further discussion.

$$dV(t) = \kappa\left(\theta - V(t)\right)dt + \varepsilon\gamma\left(V(t)\right)\left\langle dZ(t); d\rho(t)/\rho(t)\right\rangle + \varepsilon\gamma\left(V(t)\right)dZ_{k+1}(t)$$
$$= \kappa\left(\theta - V(t)\right)dt + \varepsilon\gamma\left(V(t)\right)dZ(t), \qquad (5b)$$

where W_{k+1} and Z_{k+1} are Brownian motions under $\mathbf{Q}^{k+1}$, and where the last equality in the equation for $dV(t)$ follows from Assumption 2. Without this assumption, it is clear that the SDE for $V(t)$ under $\mathbf{Q}^{k+1}$ would involve forward rates themselves, making the overall dynamics analytically intractable.

Under (3) and (4), pricing of a contingent claim C with time T payout of $C(T)$ amounts to evaluating the $\mathbf{Q}$-expectation

$$C(t) = B(t)E_t^{\mathbf{Q}}\left(C(T)/B(T)\right)$$

where $C(T)$ can depend on the entire paths of W and Z on $[0,T]$. In practice this computation is done by Monte Carlo simulation; Appendix A outlines and tests a possible algorithm.

Finally, a word about what parameterization of the function $\gamma(x)$ is most useful in applications. Scale invariance[6] alone suggests that we should use a power function (satisfying $\gamma(xy) = \gamma(x)\gamma(y)$), then the second equation becomes simply

$$dV^*(t) = \kappa\left(\theta^* - V^*(t)\right)dt + \varepsilon^*\gamma\left(V^*(t)\right)dZ(t), \quad \varepsilon^* \equiv \varepsilon q/\gamma(q).$$

In other words, if we set $\gamma(x) = x^p$ for some scalar p, then a re-scaling of the variance level can be accounted for solely by scaling the model parameters λ_k, θ, and ε. In terms of selecting a particular value of the power p, we note that application of the usual Feller boundary tests (see e.g. Karlin and Taylor (1981)) reveals that the resulting process for V can reach 0 for $p < 0.5$ (and can reach zero for $p = 0.5$ if $2\kappa\theta > \varepsilon^2$) and becomes explosive for $p > 1$. In most of our numerical work, we therefore focus on the "safe" value of $p = \frac{3}{4}$; the exact choice, however, is non-critical as the volatility smile shapes implied by the model are nearly independent of the choice of p, at least as long as $0.5 < p \le 1$. In any case, to ensure maximum flexibility, our theoretical treatment in the next two sections allows $\gamma(x)$ to be a general function.

3. Caplet Pricing Formulas.

For the model (3)-(4) to be useful in practice, it must provide for fast pricing of caps and swaptions to facilitate efficient calibration of market parameters. This section will develop asymptotic expansions for cap prices; swaptions are treated in Section 4. The accuracy of the expansions are tested in Section 5.

[6] In the sense that re-scaling V with a constant positive factor leaves (3) unchanged after appropriate scaling of the parameters $_k$, , and .

Recall that an interest rate cap with strike X is composed of so-called caplets, each paying at time T_{k+1} the amount $\delta_k \left(F_k(T_k) - X\right)^+$, with the notation $x^+ = \max(x, 0)$. By standard theory, the time t arbitrage-free price C of the caplet is given by

$$C(t) = P(t,T_{k+1})\delta_k E_t^{Q_{k+1}}\left[\left(F_k(T_k) - X\right)^+\right] \equiv P(t,T_{k+1})\delta_k G\left(t, F_k(t), V(t)\right) . \qquad (6)$$

For the purpose of pricing European options, the term $\lambda_k(t)^T dW_{k+1}(t)$ in (5a) can be replaced by $\| \lambda_k(t) \| dY(t)$, where $Y(t)$ is a one-dimensional Brownian motion uncorrelated to $Z(t)$. From the Feynman-Kac theorem (e.g. Duffie (1996)), we thereby find that the function $G(t,F,V)$ in (6) must satisfy the PDE

$$\frac{\partial G}{\partial t} + \kappa(\theta - V)\frac{\partial G}{\partial V} + \tfrac{1}{2}\varepsilon^2\gamma(V)^2\frac{\partial^2 G}{\partial V^2} + \tfrac{1}{2}\phi(F_k)^2 V \| \lambda_k(t) \|^2 \frac{\partial^2 G}{\partial F_k^2} = 0 ,$$

subject to the terminal boundary condition $G(T_k, F_k, V) = (F_k - X)^+$. Notice the absence of a mixed derivative $\partial^2 G / \partial F_k \partial V$, a consequence of our Assumption 2.

To develop asymptotic solutions to the PDE for G, we proceed in two steps, first letting V be constant, and then later allowing V to become stochastic. For notational convenience, throughout the rest of this section we drop the subscript k, such that F_k becomes F, $\| \lambda_k \|$ becomes λ (a scalar), and T_k becomes T. With this convention, the PDE to be solved is thus

$$\frac{\partial G}{\partial t} + \kappa(\theta - V)\frac{\partial G}{\partial V} + \tfrac{1}{2}\varepsilon^2\gamma(V)^2\frac{\partial^2 G}{\partial V^2} + \tfrac{1}{2}\phi(F)^2 V \lambda(t)^2 \frac{\partial^2 G}{\partial F^2} = 0 , \text{ s.t. } G(T,F,V) = (F - X)^+ \quad (7)$$

3.1. Constant *V*.

Assume for the remainder of this sub-section that V is constant; with no loss of generality, let us simply set $V = 1$. Let us also, for now, assume that $\lambda(t)^2$ is a constant, c. The PDE (7) now becomes

$$\frac{\partial G}{\partial t} + \tfrac{1}{2}\phi(F)^2 c\frac{\partial^2 G}{\partial F^2} = 0 , \text{ s.t. } G(T,F) = (F - X)^+ . \qquad (8)$$

While a closed-form solution to (8) is known for special cases of $\phi(F)$ -- most notably for the case of displaced power-type functions $\phi(F) = (aF + b)^\alpha$ -- such a solution is not possible in the general case. Instead, we here wish to develop an asymptotic series solution to (8). One approach to obtain such a series is to rely on singular perturbation techniques, as in Hagan and Woodward (1999). However, we generally find that a small-time expansion around the known solution for $\phi(F) = F$ leads to more robust results, and consequently we will present this method here.

Proposition I.

Let $\tau = T - t$. *An asymptotic expansion for the solution to* (8) *is* $G(t,F) = g(t,F;c)$, *where*

$$g(t,F;c) = F\Phi(d_+) - X\Phi(d_-)\,,\ d_\pm = \frac{\ln(F/X) \pm \frac{1}{2}\Omega(t,F,c)^2}{\Omega(t,F,c)}\,, \tag{9}$$

where Φ *is the cumulative Gaussian distribution function, and*

$$\Omega(t,F,c) = \Omega_0(F)c^{1/2}\tau^{1/2} + \Omega_1(F)c^{3/2}\tau^{3/2} + O(\tau^{5/2})\,, \tag{10}$$

$$\Omega_0(F) = \frac{\ln(F/X)}{\int_X^F \jmath(u)^{-1}du}\,,$$

$$\Omega_1(F) = -\frac{\Omega_0(F)}{\left(\int_X^F \jmath(u)^{-1}du\right)^2}\ln\left(\Omega_0(F)\left(\frac{FX}{\jmath(F)\jmath(X)}\right)^{1/2}\right).$$

Proof:

The result in Proposition I represents an expansion around the volatility term in the classical Black (1976) formula, arising in the special case $\jmath(F) = F$. We notice that the quantity $\Omega(t,F,c)/\sqrt{\tau}$ thus represents the *implied Black-Scholes volatility*. We must obviously require $\Omega(t,F,c) \sim \tau^{1/2}$ as $\tau \to 0$, and so we seek a small τ solution of the form

$$\Omega(t,F,c) = \sum_{i\geq 0} c^{i+1/2}\tau^{i+1/2}\Omega_i(F)\ . \tag{11}$$

Notice that (11) omits all integer powers of τ -- it turns out that the weights on terms of the form τ^i are all identically 0[7]. Substituting (11) into (8) and matching terms of order $\tau^{1/2}$ gives

$$F^2\Omega_0^2 = \jmath^2(F)\left(1 - \frac{\Omega_0'}{\Omega_0}F\ln(F/X)\right)^2\ , \tag{12}$$

where primes denote differentiation with respect to F. Taking square-roots of the above equation and rearranging leads to two first-order ordinary differential equations of the

[7] The proof of this statement is simple and follows directly from the structure of the resulting homogenous differential equations for these terms.

Bernoulli type. Solving (12) subject to the boundary condition that the limit of Ω_0 must be finite for $F \to X$ (and discarding the negative solution) leads to the result for Ω_0 in Proposition I.

Progressing now to the $O(t^{3/2})$ term in (10), we get

$$2F\Omega_1 = \tfrac{1}{2}\jmath^2(F)\left\{F\Omega_0'' + \Omega_0'\right\} - \jmath(F)F\ln(F/X)\left(\frac{\Omega_1'}{\Omega_1} - \frac{\Omega_1\Omega_0'}{\Omega_0^2}\right),$$

which again leads to a Bernoulli-type ODE, the explicit solution of which is as listed in the Proposition. Again, we have ensured that the limit for $F \to X$ is finite.

While developed specifically for small t, experience shows that the result in Proposition I is often remarkably accurate for long option maturities. As an illustration, the figure below tests the expansion on the strike dependency of 10-year implied caplet volatility for two CEV models and a bounded exponential model of the type $\jmath(x) = x(1 + ae^{-bx})$, $a, b > 0$. Specifically, the figure below compares the volatility skews generated by the expansion in Proposition I against either closed-form results or, in the case of the exponential model, a finely spaced Crank-Nicholson finite difference scheme.

Figure 2 Here

The expansion is extremely precise for the CEV models and also does an excellent job on the exponential model, particularly in light of the fairly extreme parameters of the latter model. Incidentally, exponential models of the type in Figure 1 are often useful in generating volatility skews that flatten out faster than CEV skews for high strikes.

As mentioned in the proof, the result in Proposition I represents an expansion around the Black-Scholes solution with the term $\Omega/\sqrt{T-t}$ conveniently representing the implied Black-Scholes volatility σ_{imp}, a quantity that is often quoted in interest rate cap markets. By equating (9) to a given market price, we can also introduce a constant scalar λ_{imp} -- the implied "skew volatility" -- as the solution to $g(t; \lambda_{imp}^2) = G(t)$, where $G(t)$ is a known, market-quoted caplet price. From Proposition I, λ_{imp} and σ_{imp} are related as

$$\sigma_{imp} \approx \Omega_0(F)\lambda_{imp} + \Omega_1(F)\lambda_{imp}^3(T-t) \ . \tag{13}$$

For a given value of σ_{imp}, this cubic equation for λ_{imp} can easily be solved in closed form, a fact that can often be exploited to speed up the calibration of the model.

Before proceeding, we point out that for many instances of $\jmath$ it is possible to further improve the accuracy of the asymptotic expansion in Proposition I by expanding

around either a displaced log-normal diffusion or a Gaussian diffusion, rather than geometric Brownian motion. The resulting series are available from the authors upon request. For the purposes of this paper, however, the result in Proposition I is easily accurate enough.

Consider now the case where $l(t)^2$ is no longer constant in time, but we still have $V = 1$. The pricing PDE then becomes

$$\frac{\partial G}{\partial t} + \tfrac{1}{2}\boldsymbol{j}(F)^2 l(t)^2 \frac{\partial^2 G}{\partial F^2} = 0 \text{ , s.t. } G(T,F) = (F - X)^+ . \tag{14}$$

Proposition II

An asymptotic solution to (14) *is given by*

$$G(t,F) = g\left(t, F; (T-t)^{-1}\int_t^T l(u)^2 du\right) ,$$

where the function g is defined in Proposition I.

Proof:

Performing a standard deterministic time-change (as in Andersen and Andreasen (2000a), Lemma 1) we find that Proposition I holds unchanged for (14), provided that we set

$$c\boldsymbol{t} = \int_t^T l(u)^2 du \ .$$

This is also easily seen from the structure of the pricing PDE. The result in Proposition II follows directly from this observation.□

3.2 Stochastic *V*.

We now let V follow (3), and seek an asymptotic solution to the full PDE (7). From Assumption 2, an adaptation of a decomposition result first proposed in Hull and White (1987) allows us to write an asymptotic solution to (7) as

$$G(t,F,V) = E_t\left[g\left(t,F;(T-t)^{-1}U(T)\right)\right],\ U(T) \equiv \int_t^T l(u)^2 V(u) du \tag{15}$$

where we understand that the expectation is with respect to the $\mathbf{Q}^{k+1}$ measure. The density of $U(T)$ is difficult to generate and generally would require transform inversion. It is easy, however, to establish the first moment of $U(T)$:

$$E_t\left[U(T)\right] = \int_t^T \lambda(u)^2 E_t\left[V(u)\right]du = \int_t^T \lambda(u)^2\left(\theta + (V(t)-\theta)e^{-\kappa(u-t)}\right)du \equiv m_U\left(t, V(t)\right) \quad (16)$$

where the second equation follows from standard results (and is subject to some standard regularity conditions on γ). With a Taylor expansion of (15) in mind, we now seek to establish higher moments of $U(T)$. Following the ideas in Lewis (2000), let us introduce the Laplace transform $E_t\left[e^{-sU(T)}\right] \equiv L\left(t, V(t); s\right)$ where s is the complex-valued argument to the Laplace-transform. L satisfies the PDE

$$\frac{\partial L}{\partial t} + \kappa(\theta - V)\frac{\partial L}{\partial V} + \tfrac{1}{2}\varepsilon^2\gamma(V)^2\frac{\partial^2 L}{\partial V^2} - s\lambda(t)^2 VL = 0\,, \text{ s.t. } L(T, V; s) = 1\,. \quad (17)$$

We now introduce a centered transform $l(t, V; s)$ by

$$l(t, V; s) = E_t\left[e^{-s\left[U(T) - m_U(t,V)\right]}\right] \Rightarrow L(t, V; s) = l(t, V; s)e^{-sm_U(t,V)}\,. \quad (18)$$

Introduction of l ensures that we focus on deviations from the mean, which are small if either ε or $\tau = T - t$ are small. Insertion of (18) into (17) yields the PDE

$$\frac{\partial l}{\partial t} + \kappa(\theta - V)\frac{\partial l}{\partial V} + \tfrac{1}{2}\varepsilon^2\gamma(V)^2\left\{\frac{\partial^2 l}{\partial V^2} + ls^2 p(t)^2 - 2sp(t)\frac{\partial l}{\partial V}\right\} = 0\,, \text{ s.t. } l(T, V) = 1\,, \quad (19)$$
$$p(t) \equiv \int_t^T \lambda(u)^2 e^{-\kappa(u-t)}du\,.$$

While we could attempt a small-time expansion for l similar to that in Section 3.1, the presence of the time-dependent term $p(t)$ in (19) makes this approach somewhat cumbersome. Instead, we expand in ε^2, the squared "volatility of variance" parameter:

Lemma I:
Let $p(t)$ *be as in* (19), *and define* $\tilde{\gamma}(x) = \tfrac{1}{2}\gamma^2(x)$ *and* $h(u) = \theta + (V - \theta)e^{\kappa(t-u)}$. *An asymptotic expansion for the solution to* (19) *is given by*

$$l(t, V; s) = 1 + \varepsilon^2 l_1(t, V; s) + \varepsilon^4 l_2(t, V; s) + O(\varepsilon^6)\,,$$
$$l_1(t, V; s) = s^2 l_{1,2}(t, V)\,,$$
$$l_2(t, V; s) = s^2 l_{2,2}(t, V) + s^3 l_{2,3}(t, V) + \tfrac{1}{2}s^4\left(l_{1,2}(t, V)\right)^2\,,$$

with

$$l_{1,2}(t, V) = \int_t^T p^2(u)\tilde{\gamma}\left(h(u)\right)du\,,$$
$$l_{2,2}(t, V) = \int_t^T e^{2\kappa u}\tilde{\gamma}\left(h(u)\right)\int_u^T e^{-2\kappa v}p^2(v)\tilde{\gamma}''\left(h(v)\right)dvdu\,,$$
$$l_{2,3}(t, V) = -2\int_t^T e^{\kappa u}p(u)\tilde{\gamma}\left(h(u)\right)\int_u^T e^{-\kappa v}p^2(v)\tilde{\gamma}'\left(h(v)\right)dvdu\,.$$

Proof:
Let

$$l(t,V;s) = 1 + \sum_{i\geq 1} l_i(t,V;s)\varepsilon^{2i}.$$

Notice that odd powers on ε are not used in the expansion, as only ε^2 figures in the PDE (19). Inserting this expression into (19) and collecting terms of order ε^2 gives

$$\frac{\partial l_1}{\partial t} + \kappa(\theta - V)\frac{\partial l_1}{\partial V} + \tfrac{1}{2}s^2\rho(t)^2\gamma(V)^2 = 0, \text{ s.t. } l_1(T,V) = 0.$$

This simple PDE can be solved in closed form, yielding the expression for l_1 above. The result for l_2 is established by collecting terms of order ε^4 and proceeding as for l_1.□

While somewhat complicated in appearance, the expressions for the integrals $l_{1,2}$, $l_{2,2}$, and $l_{2,3}$ are trivial to implement on a computer. Indeed, due to the nested nature of the double integrals $l_{2,2}$ and $l_{2,3}$, all integrals can be computed in a single numerical integration loop, at negligible computational cost. In some cases of practical interest it is also possible to evaluate the integrals analytically.

Equipped with Lemma I, we can compute central moments as follows:

$$E_t\left[\left(U(T) - \mu_U(t,V)\right)^n\right] = (-1)^n \partial^n l(t,V;s)/\partial s^n\Big|_{s=0},\ n = 2,3,..$$

There are many ways to turn these moments into an option price expression. For instance, we could rely on an Edgeworth expansion or perhaps some parametric density family to express the full density of $U(T)$. Eq. (15) could then be evaluated by integration of the density-weighted expression for the function g. Alternatively, we can Taylor expand (15) and produce a closed-form asymptotic result. This latter approach results in the following asymptotic solution to the option pricing problem (7):

Proposition III:
Define $Y = \ln(F/X)$ *and* $\bar{\sigma}(t,V) = \tau^{-1}\mu_U(t,V)$. *An asymptotic expansion for the caplet price* $G(t,F,V)$ *that solves* (7) *is*

$$G(t,F,V) = g\left(t,F;\sigma^*(t,V)\right),$$

where the function g *is given in Proposition I, and to order* ε^2

$$c^*(t,V) = \overline{c}(t,V) + \alpha_0\varepsilon^2 + \alpha_1\varepsilon^2 Y^2 + O(\varepsilon^4), \tag{20}$$

or to order ε^4

$$c^*(t,V) = \overline{c}(t,V) + (\alpha_0\varepsilon^2 + \beta_0\varepsilon^4) + (\alpha_1\varepsilon^2 + \beta_1\varepsilon^4)Y^2 + \beta_2\varepsilon^4 Y^4 e^{-\Lambda\varepsilon^2 Y^2} + O(\varepsilon^6), \tag{21}$$

where the coefficients α_0, α_1, β_0, β_1, β_2 *are given in Appendix B, and where* Λ *is an arbitrary positive number. The corresponding Black-Scholes implied volatility is given by*

$$\sigma_{imp} = \Omega_0\sqrt{c^*(t,V)} + \Omega_1 c^*(t,V)^{3/2}\tau + O(\tau^2), \tag{22}$$

where Ω_0 *and* Ω_1 *are given in Proposition I.*

Proof: In Appendix B. ⌟

We notice that $\sqrt{c^*} = \lambda_{imp}$, where λ_{imp} is the implied "skew-volatility" introduced earlier. In the special case of $\phi(x) = x$, we have $\lambda_{imp} = \sigma_{imp}$ and, provided that λ is constant, the expression (20) becomes identical to the result given in Hull and White (1987) and many other sources. As we shall see in Section 5, (20) is adequate for reasonably low values of ε (or, alternatively, high values of the mean reversion parameter κ). In other situations, (20) is somewhat inaccurate, particularly (and unfortunately) for at-the-money options where $Y = 0$, in which case the higher-order result (21) should be used. (21) includes a Y^4 term and corrects both the at-the-money term and the Y^2 term. As we are here less interested in the contributions from the Y^4 term (which tend to grow uncontrollably for strikes far away from the forward), we have introduced a transcendental dampening factor $e^{-\Lambda\varepsilon^2 Y^2}$ on this term. The choice of Λ is non-critical, as long as it is big enough to prevent the Y^4 term from causing problems for high values of $|Y|$. In most applications, a number between 1 and 10 is sufficient. We note that it is comparatively straightforward, but tedious, to continue the expansions to order ε^6, should additional precision be required.

4. Swaption Pricing Formulas.

To facilitate calibration to swaptions, this section will discuss how to apply the expansion results for caps and floors to approximate European swaption prices for the process dynamics (3)-(4). The development of these formulas closely follow the approach taken in Andersen and Andreasen (2000a), and certain steps will be omitted for brevity.

Consider a European payer swaption maturing at some date T_s, $s \in \{1,2,...,K\}$. The swaption gives the holder the right to pay fixed cashflows $\alpha\delta_{k-1} > 0$ at T_k, for $k = s+1, s+2,...,e$ in exchange for Libor (paid in arrears) on a \$1 notional. α is thus the annual coupon rate, and T_s and T_e are the start- and end-dates of the underlying

swap, respectively; clearly we require $T_{K+1} \geq T_e > T_s$. At maturity T_s the value of the payer swaption S is, by definition,

$$S(T_s) = \left(\sum_{k=s}^{e-1} P(T_s, T_{k+1}) \delta_k [F_k(T_s) - \alpha] \right)^+ = B^S(T_s) \left(R(T_s) - \alpha \right)^+, \tag{23}$$
$$B^S(t) \equiv \sum_{k=s}^{e-1} \delta_k P(t, T_{k+1}), \quad R(t) \equiv \frac{P(t, T_s) - P(t, T_e)}{B^S(t)},$$

where the second equality for $S(T_s)$ requires a few straightforward manipulations. $B^S(t)$ in (23) is a strictly positive process and can thus be used as a pricing numeraire. The probability measure induced by B^S is denoted $\mathbf{Q}^S$, and is often referred to as the *(forward) swap measure* for the swaption S. By standard theory, absence of arbitrage implies that the *par-rate* $R(t)$ is a martingale under $\mathbf{Q}^S$. Assuming that the yield curve dynamics are governed by the stochastic volatility model (3)-(4), an application of Ito's lemma yields,

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

where W^S is an m-dimensional Brownian motion under $\mathbf{Q}^S$ and $\partial R(t) / \partial F_j$ is easily computed analytically. The stochastic differential equation above for $R(t)$ is not analytically tractable. To proceed, we notice that for most reasonable shifts of the forward curve, $\partial R(t) / \partial F_j$ normally varies little with time and the state of interest rates. In many cases, it is often also reasonable to assume that the ratio $\psi\left(F_j(t)\right) / \psi\left(R(t)\right)$ is close to constant. The quality of this assumption obviously depends on the specific form of ψ, but seems in practice to hold very well for many useful functions, particularly those that are close to the power-type (CEV) specification (see Andersen and Andreasen (2000a) for a variety of tests). In total, we suggest the following approximation of the dynamics of R in the model (3)-(4):

$$dR(u) \approx \psi\left(R(u)\right) \sqrt{V(u)} \sum_{j=s}^{e-1} w_j(t) \lambda_j^{\mathrm{T}}(u) dW^S(u), \ t \leq u < T_s, \tag{24}$$
$$w_j(t) = \frac{\partial R(t)}{\partial F_j(t)} \frac{\psi\left(F_j(t)\right)}{\psi\left(R(t)\right)}.$$

With the approximation (24), the SDE for R under $\mathbf{Q}^S$ takes exactly the same form as the forward rate SDE's in $\mathbf{Q}^{k+1}$; see Section 2. As such, the swaption pricing problem becomes identical (after substitution of numeraires) to the caplet pricing problem discussed in detail in Section 3. Specifically, we can write the swaption price as

$$S(t) = B^S(t)H\big(t, R(t), V(t)\big),$$

where $H(u,R,V), t \le u \le T$, solves

$$\frac{\partial H}{\partial u} + \kappa(\theta - V)\frac{\partial H}{\partial V} + \tfrac{1}{2}\varepsilon^2\gamma(V)^2\frac{\partial^2 H}{\partial V^2} + \tfrac{1}{2}\rho(R)^2 V^2 \lambda_R(u)^2\frac{\partial^2 H}{\partial R^2} = 0,\ H(T_s,R,V) = (R-\alpha)^+, \tag{25}$$

$$\lambda_R(u) \equiv \left\| \sum_{j=s}^{e-1} w_j(t)\lambda_j(u) \right\|.$$

(25) is, apart from notation, identical to (7) whereby all expansion results of Section 3 can be applied directly.

5. Numerical Results.

In this section we will investigate the numerical properties of the suggested framework and test the precision of the asymptotic expansions for cap and swaption pricing.

5.1. Cap pricing.

As our first test case, we set $F(0) = 6\%$ for all forward maturities, $\lambda(t)$ to a constant 20%, $\gamma(x) = x^{3/4}$, and, as we first wish to focus exclusively on the stochastic volatility expansion, choose a log-normal base model, $\rho(x) = x$. We denote this choice of parameters **Test A**, and shall proceed to test the performance of the expansions (20) and (21) as functions of maturity T, mean reversion speed κ, and volatility of variance ε. For ease of reference we denote the combination of (20) and (22) **expansion O2** (as it is accurate to second order); the combination of (21) and (22) is denoted **expansion O4**.

In Figure 2 we have fixed $T = 1.5$ and $\kappa = 1$ (implying a half-life of volatility shocks of 1 year) and demonstrate the performance of the expansion solution (22) for σ_{imp}, as a function of option moneyness and ε. The panels of the figure use quite extreme values of ε (150% and 200%, respectively) to generate visible differences between the ADI and O4 smiles. Notice the difference in scales for the implied volatility used in the two panels: as ε is increased the magnitude of the volatility smile implied by the model increases, as expected.

Figure 3 Here

As the figure shows, the expansion precision is generally excellent across a broad range of strikes, with the higher-order series in Eq. (21) outperforming the one in Eq. (20). The precision of both expansions decreases for large values of ε, as one would expect,

although the breakdown in precision is quite gradual. For instance, for $\varepsilon = 200\%$ the at-the-money implied volatility error for the O4 expansion is 0.00013 (or 0.13%), easily within the typical at-the-money bid-offer spread of around 0.5% - 1.0% (bid-offer spreads are significantly higher for options with strikes away from at-the-money). For $\varepsilon = 150\%$ the at-the-money error of the O4 expansion is 0.02%.

In Figure 4 we have fixed ε at 150% and investigate the effects of varying time to maturity (T).

Figure 4 Here

As expected, the effects of stochastic volatility on the implied volatility smile dies out over time, a consequence of mean reversion. Notice from the figure that the performance of the stochastic volatility expansion holds up well, even for long maturities. Notice that the errors of both expansions eventually decline as T gets very large (this is evident in panel B for expansion O4, but a similar effect holds for expansion O2). While figure 4 has fixed ε at 150%, the results for other values of ε are qualitatively similar.

As we demonstrated earlier, increasing the vol-of-variance parameter ε increases the steepness of the implied volatility smile for any fixed maturity T. A similar effect can be obtained by keeping ε fixed and decreasing the mean reversion parameter κ. In Figure 5, panel A, we demonstrate the effects on the 1.5-year caplet smile of moving κ from 1 to the extreme value of 0.

Figure 5 Here

The O4 series expansion generally shows significantly less sensitivity to decreases in κ than does the O2 expansion, particularly for the important case of at-the-money options. Figure 5, panel B shows in more detail the series performances for at-the-money options at various mean reversion speeds. The graph has been done for 2-year caplets; other maturities display qualitatively similar shapes. One caveat: for maturities beyond, say, 4-5 years, usage of κ's close to zero generally becomes questionable unless ε is small (say, less than 50%), as the variability of volatilities over long horizons becomes excessive and in some cases can make the distribution of $V(T)$ degenerate for large T. In such situations it becomes increasingly difficult to make numerical methods (including finite difference schemes and Monte Carlo simulations) converge, and it might be necessary to add further terms to the asymptotic expansions.

Before we move on to models with more complicated skew functions j, we consider briefly the role of the truncation factor Λ. For the parameters and strike-ranges considered so far, the choice of Λ is non-critical – our selection of $\Lambda = 1$ for the numerical experiments was of little consequence for the displayed results and we could equally well have set $\Lambda = 0$. However, for very large values of $|\ln(F/K)|$ -- that is, for

options very deeply in-the-money or out-of-the-money -- the order Y^4 term in expansion O4 can become problematic, particularly in situations where the model parameters are extreme (high ε, low κ). Figure 6 demonstrates such a situation. Notice the effect of the defensive term $e^{-\Lambda\varepsilon^2 Y^2}$ which overall makes the expansion more robust without sacrificing much accuracy. (We should point out that the monetary impact of the break-down of the expansion for large $|Y|$ is limited as the sensitivity of caplet prices to volatility is very low in this region. As a consequence, bid-offer spreads in implied volatility terms tend to be very high for strikes far away from the forward).

Figure 6 Here

So far we have confined our investigation to the case where $\beta(x) = x$. In these cases, our asymptotic expansions approximate the volatility smile as near-parabolic in $\ln(F(0)/X)$. In particular, in this model the volatility smile generated by the expansion is always symmetric in moneyness $F(0)/X$ for X close to $F(0)$, a consequence of having zero correlation between the forward rate and variance processes. To generate a non-symmetric smile, we now move to **Test B** where we change our base skew assumption and set $\beta(x) = x^{0.6}$. The rest of the Test A parameters are kept, although we now renormalize λ to 6.49% $(= 20\% \cdot F(0)^{1-0.6})$ to retain approximately the same volatility levels as in test A. Figure 7 shows the 1.5-year volatility smile for Test B with $\kappa = 1$ and ε equal to 150% and 200%, respectively.

Figure 7 Here

We note that in Figure 7, the parabolic smile effect of stochastic volatility is now superimposed on a non-flat base skew, effectively generating an asymmetric "smirk", a shape that is empirically observable in a range of markets. As for the performance of the expansions, the results are both qualitatively and quantitatively similar to Figure 3 (which use the same *V*-process parameters, but $\beta(x) = x$). This is entirely to be expected given the high precision of the expansion in Proposition I-II (see Figure 2) and the orthogonality of the forward rate and variance processes. Indeed, in virtually all practical applications, the ε-expansion in Proposition III is more critical than the small-time skew expansion in Proposition I-II. Having already spent considerable time testing the former expansion, we shall not repeat the tests for the Test B setting, but just point out that the performance of the expansions O2 and O4 in Test B for various levels of T, ε, κ is consistent with the results shown for Test A in Figures 3-6. Instead, we show in Figure 8 an example of the full caplet volatility smile surface that can be generated for the settings in Test B. Notice that the intial "smirk" gradually turns into a skew as the caplet maturity is increased. The speed at which this transition takes place is, as we have seen, a function of the volatility of variance ε and the mean reversion speed κ.

Figure 8 Here

5.2. Swaption Pricing.

We now turn to the test of our pricing formulas applied to European swaptions. From the results of Section 4, we notice that the accuracy of the swaption pricing formulas depend on both the precision of the expansions in Propositions I and II, as well as the approximations made to arrive at the simplified SDE (25). The asymptotic expansions have been extensively tested in Section 5.1., and the validity of the steps leading to (25) have been examined at length in Andersen and Andreasen (2000a) as well as Hull and White (2000), among several others. In the interest of brevity, we shall therefore limit ourselves to a few representative tests, all set within the following, fairly realistic, two-factor scenario:

Test C: Forward curve: $\delta = 0.5$, $F_k(0) = 0.04 + 0.00075 \cdot k$. Skew function: $\beta(x) = \sqrt{x}$. Stochastic volatility dynamics: $V(0) = \theta = 1, \kappa = 1, \gamma(x) = x^{3/4}$, and $\varepsilon = 1.5$. Volatility term structure: $\lambda_k(T_j) = \left(0.015 + 0.025e^{-0.05(k-j)}, 0.01 - 0.05e^{-0.1(k-j)}\right)^{\mathbf{T}}$, with $\lambda_k(t)$ piecewise flat between dates in the tenor structure.

The above scenario involves an upward-sloping linear yield curve with $F_0(0) = 4\%$ and $F_{40}(0) = 7\%$, and a two-factor time-stationary volatility term structure. The net instantaneous forward volatility $\| \lambda_k(t) \| = \sqrt{\lambda_k^1(t)^2 + \lambda_k^2(t)^2}$ in Test C is downward-sloping in $T_k - t$ and will generate implied swaption volatilities that decrease both in swap tenor and swaption maturity, consistent with typical market conditions[8]. In log-normal terms, $\| \lambda_k(0) \|$ starts out at around 28% for $k = 0$ and decays to around 8% for $k = 40$. The instantaneous correlation between F_0 and F_{40} is 32%.

In Table 1, we list for Scenario C the prices of at-the-money swaptions with maturities (T_s) and swap tenors ($T_e - T_s$) ranging between 1 and 10 years. Prices have been computed both by Monte Carlo simulations and the fourth-order expansion (21)-(22), applied as discussed in Section 4. The accuracy of the expansion results is excellent, with all errors being less than 1% of the simulated price (and all less than 0.15% in terms of implied volatility). Indeed, for the 50,000 antithetic paths used in the table, the differences between the expansion and Monte Carlo results are mostly not statistically significant. Recall also (see Appendix A) that the Monte Carlo simulated results themselves are subject to a discretization error, although it is likely to be small here.

[8] This phenomenon is consistent with mean reversion in interest rate movements. While not used in our numerical example, we note that the stochastic volatility component of our model can produce downward-sloping implied volatilities, even for constant $\quad_k$, if $V(0) >$.

Table 1: Prices of At-the-Money Swaptions in Scenario C

T_s	$T_e - T_s$	R	MC (SD)	Expansion O4	Error	Error (%)	Error (Vol)
1	1	4.19%	35.67 (0.15)	35.45	0.23	0.64%	0.15%
1	5	4.47%	134.11 (0.53)	133.29	0.82	0.61%	0.11%
1	10	4.80%	204.47 (0.78)	203.14	1.33	0.65%	0.09%
5	1	4.79%	57.62 (0.24)	57.15	0.47	0.82%	0.14%
5	5	5.07%	223.73 (0.85)	222.17	1.56	0.70%	0.10%
5	10	5.40%	350.55 (1.25)	348.77	1.78	0.51%	0.06%
10	1	5.54%	58.27 (0.22)	57.89	0.37	0.64%	0.09%

Notes: The table above shows at-the-money (in the sense that the swap coupon equals the par rate, $R = \alpha$) payer swaption prices in model scenario C, as produced by the expansion (21)-(22) ("Expansion O4") and Monte Carlo simulation ("MC"). Numbers in parentheses refer to sample standard deviations ("S.D") as generated by 50,000 antithetic paths (for a total of 100,000 separate paths). "Error" refers to the difference between the prices generated by expansion and Monte Carlo simulation; "Error (%)" is the error as a percentage of the Monte Carlo result; and "Error (Vol)" is the error expressed in terms of implied (Black) volatility. Prices and errors are reported in basis points (1/10,000). The table contains results for swaption maturities (T_s) and swap tenors ($T_e - T_s$) ranging from 1 to 10 years; for reference, the third column lists the par rate (R) of each swaption. In the application of the expansion in Eq. (21), the cut-off parameter Λ was set to 1. The Monte Carlo results reported in the graph were computed with 50,000 antithetic paths (a total of 100,000 separate paths) and the simulation algorithm (A.1)-(A.2) in Appendix A with a constant time step of $\Delta = \frac{1}{8}$.

To examine the performance of the expansion results for swap coupons away from the par rate we now freeze the tenor to $T_e - T_s = 5$ years (results are comparable for other tenors) and examine the full volatility smile for various swaption maturities; see Figure 9 below. It is clear from the figure that the expansion results work well even for swaptions that are far from at-the-money.

Figure 9 Here

7. Parameter Estimation and Model Extensions.

The way we envision that the proposed framework typically would be parameterized involves calibrating the l_k vectors to observed prices of swaptions and caps (see Andersen and Andreasen (2001) and Sidenius (1999) for details) for fixed exogenous specification of $\gamma(x)$ and the parameters of the V process, primarily ε and κ . The parameterization of $\gamma(x)$ and the stochastic volatility parameters could be

obtained from observations of the dynamics of implied volatilities as well as the market skew/smile and its dependence of option maturity. (Note that the V process parameters must be estimated in the spot measure, not the statistical measure. As such, estimation directly from empirical forward rate dynamics is not possible, and we instead must look at option prices and their implied volatilities). We refer to Joshi and Rebonato (2001) for some observations about estimating the parameters from empirical observations of smile and skew behavior, and here just demonstrate how, in principle, information about the stochastic volatility process, as well as the skew function, can be extracted from dynamics of at-the-money implied volatility data.

Let us define $V^t_{imp}(t) = V^t_{imp}(t,F,V)$ as the time t at-the-money implied log-normal variance (=volatility squared) of a caplet maturing at time $t+\tau$. F is the forward rate underlying the caplet, and V is, as before, the state of the variance multiplier process. Applying Ito's lemma to the process (3)-(4), we get that percentage changes in implied variance satisfy

$$\begin{aligned} dV^t_{imp}(t) / V^t_{imp}(t) = \ldots dt + V^t_{imp}(t)^{-1} \frac{\partial V^t_{imp}}{\partial F} \jmath\left(F(t)\right) \lambda(t) dW(t) \\ + V^t_{imp}(t)^{-1} \frac{\partial V^t_{imp}}{\partial V} \varepsilon\psi\left(V(t)\right) dZ(t), \end{aligned} \tag{26}$$

where the drift term is measure specific. We notice that movements in implied volatilities originate from two sources: a term from movements of forward rates and a term from the V process. The first term only exists in models with non-flat base skews and simply reflects the fact that log-normal volatilities are functions of forward rate levels[9] when $\jmath(x) \neq x$ (if $\jmath(x) = x$, $\partial V^t_{imp} / \partial F = 0$). The second term exists as long as $\varepsilon > 0$ and will, due to mean reversion, affect short-dated implied volatilities more than long-dated ones. For sufficiently long-dated options, the effect of the V process will be negligible and the total volatility of implied variances will be determined solely by the skew function and $\lambda(t)$.

From (26), we see that the total volatility of implied variance is

$$\sigma_V(t;\tau) = V^t_{imp}(t)^{-1} \sqrt{\left(\frac{\partial V^t_{imp}\left(t,F(t),V(t)\right)}{\partial F} \jmath\left(F(t)\right) || \lambda(t) || \right)^2 + \left(\frac{\partial V^t_{imp}\left(t,F(t),V(t)\right)}{\partial V} \varepsilon\psi\left(V(t)\right) \right)^2}$$

an expression that for most parameterizations is fairly insensitive to the levels of V and F. The derivatives in this expression can be computed directly from the pricing formulas given in Section 5. For $\varepsilon = 130\%$, $\kappa = 3$, $\psi(x) = x^{3/4}$, $F(0) = 6\%$, and $\lambda\jmath\left(F(0)\right) = 1.2\%$,

[9] For instance, in a CEV model with $\jmath(x) = x^p$, $0 < p < 1$, implied at-the-money volatilities increase (decrease) when forward rates move up (down).

the figure below shows the time 0 volatility of implied variance as a function of t for different parameterizations of j . The figure also contains empirical results, as computed from 2 years of weekly Bloomberg data on implied volatilities of options on 1-year swaps.

Figure 10 Here

While the comparison between the empirical data and the theoretical results here is only suggestive, we do note that the model at least captures the important characteristics of empirical implied volatility dynamics: a) for short maturities, the volatility of implied variances decays rapidly as a function of t ; b) in the presence of a base skew in the model, the volatility of implied variances approaches a non-zero asymptotic value. In principle, one can use the level of the asymptote to parameterize the skew function j , although it is an outstanding question how consistent such a parameterization would be with parameterizations based on either the market skew or the empirical relationship between forwards and implied volatilities (see Figure 1). We leave this question for future work, but point out that it is possible to adjust the asymptotes in Figure 10 by a mechanism other than a DVF function, if needed. Specifically, in the base model (3)-(4), consider writing

$$\begin{aligned} V(t) &= V_1(t) + V_2(t)\,, \\ dV_1(t) &= \kappa_1\left(\theta_1 - V_1(t)\right)dt + \varepsilon_1\gamma_1\left(V_1(t)\right)dZ_1(t)\,, \\ dV_2(t) &= \varepsilon_2\gamma_2\left(V_2(t)\right)dZ_2(t)\,, \end{aligned} \tag{27}$$

where Z_1 and Z_2 are independent of each other and of the Brownian motions driving forward rates. In other words, the variance process now has a component (V_2) that does not mean revert and consequently affects the volatility of implied variances even for long maturities. Now, in equation (15) we get

$$U(T) = \int_t^T \lambda(u)^2 V_1(u)du + \int_t^T \lambda(u)^2 V_2(u)du \equiv U_1(T) + U_2(T)\,,$$

where U_1 and U_2 are statistically independent. Due to independence, we have

$$E\left[\left(U(T) - E[U(T)]\right)^n\right] = E\left[\left(U_1(T) - E[U_1(T)]\right)^n\right] + E\left[\left(U_2(T) - E[U_2(T)]\right)^n\right]$$

where each of the terms on the right hand side can be determined using the methods in Section 3. This again allows us to carry out an asymptotic expansion for cap and swaption prices with relative ease (the result is a simple extension to the results already given and is omitted). As ε_2 is likely to be small (less than 50%), the asymptotic

expansions will generally have no problems handling the lack of mean reversion in V_2, even for long maturities.

We point out that extensions of the model along the lines of (27) can be generalized to

$$V(t) = \sum_i c_i(t) V_i(t)$$

where the c_i are deterministic functions and where the V_i are independent mean-reverting processes. Such a formulation could be useful if one suspects that the variance process has multiple mean reversion time scales. Again, the asymptotic expansions for caps and swaptions can be carried out with no complications, as sketched above.

The empirical data in Figure 10 only covers a specific tenor of rate (1 year). In practice, the behavior of volatilities of implied variances is rate tenor dependent, as shown in Figure 11:

Figure 11 Here

In particular, note from the figure above that the effective "volatility of variance" parameter ε decreases with increasing tenor. This phenomenon is likely to be a diversification effect, reflecting the fact that the variance processes of the forward rates underlying a particular swap rate in practice are not, unlike in our model, perfectly correlated. In effect, our model corresponds only to the first principal component of the movements of the instantaneous volatility surface. To capture the tenor effect in our model, we would need to extend the number of factors driving the variance process, most generally by equipping each forward rate F_k with a unique mean-reverting variance process $V_k(t)$, $k = 1,2,...$ In addition, we would need to specify a correlation structure between the Brownian motions driving the various V_k-processes. We note that such an extension will cause no problems when deriving expressions for caplet pricing formulas (our results carry over unchanged) but makes the derivation of a closed-form swaption pricing formula more difficult. In addition, estimation of the many parameters in the resulting model will be challenging, although it should be possible (by a principal components analysis of movements of implied volatilities, for instance) to reduce the dimension of the estimation problems somewhat through simplifying assumptions. We leave further exploration of these issues to future research.

8. Conclusions.

This paper has developed a straightforward yet flexible framework for the incorporation of mean-reverting stochastic volatility behavior into DVF-extended Libor Market models. The proposed framework allows for arbitrary base skews and

accomodates non-parametric volatility term structure specification. In addition, the proposed model dynamics are amenable to analytical methods, and a key contribution of this paper has been the establishment of robust and accurate expansions for plain-vanilla option prices. The primary focus of the paper has been on developing the basic model framework and analyzing its analytical and numerical properties, although we have included a brief discussion of certain empirical issues along with suggestions for how the basic framework can be extended to provide an increasingly realistic model of rate dynamics. However, much interesting empirical research remains to be done; hopefully the algorithms and formulas developed in this paper will be of use in this work. Another line of investigation that we postpone for future research involves the investigation of the effects of stochastic volatility on prices of exotic structures, such as Bermudan swaptions and trigger swaps (see e.g. Andersen and Broadie (2001) for a discussion about the pricing of Bermudan swaptions in a LM model setting). Finally, it would be of interest to examine LM models that merge the approach taken in this paper with the jump model of Glasserman and Kou (1999). The resulting model would likely be challenging to parameterize but would provide a very rich framework for empirical and theoretical work.

References:

Andersen, L. and Andreasen, J. (2000a). "Volatility Skews and Extension of the Libor Market Model," *Applied Mathematical Finance*, 7, 1-32.

Andersen, L. and Andreasen, J. (2000b). "Jump Diffusion Processes: Volatility Smile Fitting and Numerical Methods for Option Pricing," *Review of Derivatives Research*, 4, 231-262.

Andersen, L. and Andreasen, J. (2001). "Factor-Dependence of Bermudan Swaptions: Fact or Fiction?" *Journal of Financial Economics*, 62, 3-37.

Andersen, L. and M. Broadie (2001). "A Primal-Dual Algorithm for Pricing American Options," Working Paper, Columbia University and Gen Re Securities.

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

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.

Chen, R.R. and L. Scott (2001). "Stochastic Volatility and Jumps in Interest Rates: An Empirical Analysis," Working Paper, Rutgers University and Morgan Stanley.

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 S. Kou (1999). "The Term Structure of Simple Forward Rates with Jump Risk," Working Paper, Columbia University.

Hagan, P. and D. Woodward (1999). "Equivalent Black Volatilities," *Applied Mathematical Finance*, 6, 3, 147-157

Hagan, P. (2000). "Stochastic Beta Models," Working Paper, Numerix.

Heston, S. (1993). "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *The Review of Financial Studies*, 6, 2, 327-343.

Hull, J. and A. White (1987). "The Pricing of Options on Assets with Stochastic Volatilities," *The Journal Finance*, 42, 281-300.

Hull, J. and A. White (2000). "Forward Rate Volatilities, Swap Rate Volatilities, and the Implementation of the Libor Market Model," Working Paper, University of Toronto.

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

Jamshidian, F. (1999). "LIBOR Market Model with Semi-Martingales," Working Paper, NetAnalytic Ltd.

Joshi, M. and R. Rebonato (2001). "A Stochastic-Volatility, Displaced-Diffusion Extension of the LIBOR Market Model," Working Paper, Royal Bank of Scotland.

Karlin, S. and Taylor, H. (1981). *A Second Course in Stochastic Processes*, Academic Press, Boston.

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

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

Merton, R. (1976). "Option Pricing when Underlying Stock Returns are Discontinuous," *Journal of Financial Economics*, May, 125-144.

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

Sidenius, J. (2000). "Libor Market Models in Practice," *The Journal of Computational Finance*, 3, 3, 5-26.

Appendix A
Monte Carlo Simulation of (3)-(4)

To implement a Monte Carlo simulation scheme for the proposed model, we need to discretize the continuous-time dynamics in Eqs. (3)-(4) on some time-grid $\{t_i\}$, $i = 1,2,\ldots$. A variety of discretization schemes are discussed in Kloeden and Platen (1992). Andersen and Andreasen (2000a) adopt a simple log-Euler scheme for the SDE (2) and demonstrate that discretization errors are manageable; applications of higher-order schemes to the log-normal version of (2) can be found in Brotherton-Ratcliffe (1997). We shall here just list the log-Euler scheme for (4):

$$\hat{F}_k(t_{i+1}) = \hat{F}_k(t_i)\exp\left(\frac{\jmath\left(\hat{F}_k(t_i)\right)}{\hat{F}_k(t_i)}\sqrt{\hat{V}(t_i)}\,\lambda_k^T(t_i)\left[\sqrt{\hat{V}(t_i)}\Delta_i\left(\hat{\mu}_k(t_i) - \tfrac{1}{2}\frac{\jmath\left(\hat{F}_k(t_i)\right)}{\hat{F}_k(t_i)}\,\lambda_k(t_i)\right) + \tilde{\nu}_i\sqrt{\Delta_i}\right]\right), \tag{A.1}$$

$$\hat{\mu}_k(t_i) = \sum_{j=\eta(t_i)}^{k} \lambda_j(t_i)\frac{\delta_j\jmath\left(\hat{F}_j(t_i)\right)}{1+\delta_j\hat{F}_j(t_i)},$$

where $\Delta_i = t_{i+1} - t_i$, and $\tilde{\nu}_i$, $i = 1,2,..$ is an independent series of draws of a m-dimensional standard Gaussian distribution. We notice that this scheme by construction will not generate negative forward rates.

To complete (A.1), we need to discretize the V-process (3). This generally requires more care than is the case for the forward rate process, particularly if the mean reversion parameter κ is high. Indeed, a direct Euler (or log-Euler) discretization of (3) is likely to be unstable for practical time steps, with the drift term oscillating between high and low values. One way to correct this problem is to apply implicit or predictor-corrector schemes; see Kloeden and Platen (1992) for details. An alternative, which we shall discuss here, approximates (3) over $[t_i, t_{i+1}]$ with a Gaussian Ornstein-Uhlenbeck process (notice that the diffusion term is evaluated only at t_i):

$$dV(t) \approx \kappa\left(\theta - V(t)\right)dt + \varepsilon\gamma\left(V(t_i)\right)dZ(t), \quad t \in [t_i, t_{i+1}].$$

Solving this approximation analytically gives rise to the scheme

$$\hat{V}(t_{i+1}) = \theta + \left(\hat{V}(t_i) - \theta\right)e^{-\kappa(t_{i+1}-t_i)} + \tilde{z}_i\varepsilon\gamma\left(\hat{V}(t_i)\right)\sqrt{\tfrac{1}{2}\kappa^{-1}\left(1 - e^{-2\kappa(t_{i+1}-t_i)}\right)},$$

where $\tilde{z}_i$, $i = 1,2,..$ is a series of univariate Gaussian draws, independent of the $\tilde{\nu}_i$ used in (A.1). This scheme is easily verified to be stable, and its convergence order is the same as that of (A.1), namely $O(\Delta)$. We notice, however, that the scheme can result in

negative values of V, a problem can be solved either by modifying the behavior of $\hat{V}$ around the origin (for instance reflecting the process off the orgin) or by using a moment matched, log-normal approximation:

$$\hat{V}(t_{i+1}) = \left(\theta + \left(\hat{V}(t_i) - \theta\right) e^{-\kappa(t_{i+1}-t_i)}\right) e^{-\frac{1}{2}\Gamma(t_i)^2 + \Gamma(t_i)\tilde{z}_i}, \tag{A.2}$$

$$\Gamma(t)^2 = \ln\left(1 + \frac{\frac{1}{2}\varepsilon^2\gamma\left(\hat{V}(t_i)\right)^2\kappa^{-1}\left(1 - e^{-2\kappa(t_{i+1}-t_i)}\right)}{\left(\theta + \left(\hat{V}(t_i) - \theta\right)e^{-\kappa(t_{i+1}-t_i)}\right)^2}\right).$$

The table below compares at-the-money caplet prices produced by Monte Carlo simulation of (A.1)-(A.2) to the prices produced by a finely spaced ADI finite difference grid. The table demonstrates that the bias induced by the combined discretization scheme (A.1)-(A.2) is typically very low, even for relatively coarse time grids. For instance, at a time-step of 0.5, the discretization error is generally around 1% of the caplet price (or here around 0.3-0.4 basis points) which is in line with typical bid-offer spreads. At finer steps, the discretization error would typically become hard to distinguish from the usual random Monte Carlo error. In the table, this is the case for $\Delta = 0.125$ and for most of the results with $\Delta = 0.25$. Specifically, at these time-steps, the observed pricing error is generally within a few sample standard deviations, making the bias statistically insignificant. The number of trials in the table (50,000, with antithetic sampling) is probably reasonably representative, or even on the high side, of what would be used in practical applications.

The analysis above can be repeated for caplets with strikes away from at-the-money, but results are similar to those of Table A.1. We refer to Andersen and Andreasen (2000a) for further numerical results and a more thorough discussion of simulation biases in extendended LM models.

Table A.1: Errors in Monte Carlo Simulation of Caplet Prices

		$T = 1$	$T = 2$	$T = 4$	$T = 6$	$T = 10$
	ADI	21.05	27.90	35.15	38.38	39.22
$\Delta = \frac{1}{8}$	MC (S.D)	21.28 (0.09)	28.02 (0.12)	35.2 (0.15)	38.44 (0.15)	39.14 (0.13)
	Error	0.22	0.12	0.05	0.06	-0.07
	Error (%)	1.06%	0.43%	0.14%	0.15%	-0.19%
$\Delta = \frac{1}{4}$	MC (S.D)	21.39 (0.09)	28.17 (0.12)	35.22 (0.15)	38.42 (0.15)	39.04 (0.13)
	Error	0.34	0.28	0.07	0.04	-0.17
	Error (%)	1.60%	0.99%	0.20%	0.10%	-0.45%
$\Delta = \frac{1}{2}$	MC (S.D)	21.58 (0.09)	28.39 (0.12)	35.58 (0.15)	38.79 (0.15)	39.39 (0.13)
	Error	0.53	0.49	0.43	0.41	0.17
	Error (%)	2.51%	1.76%	1.23%	1.06%	0.43%

Notes: The table above shows at-the-money caplet prices in a LM model with stochastic volatility and skew function $\boldsymbol{j}(x) = \sqrt{x}$, as produced by an alternating directions implicit finite difference grid ("ADI") and Monte Carlo simulation ("MC") with the scheme (A.1)-(A.2). Numbers in parentheses refers to sample standard deviations ("S.D") as generated by 50,000 antithetic paths (for a total of 100,000 separate paths). "Error" refers to the difference between the Monte Carlo prices and the ADI results. "Error (%)" is the error as a percentage of the ADI result. All numbers except for the percentage errors are reported in basis points (1/10,000). The table contains results for caplet maturities (T) ranging from 1 to 10 years, and for time-discretizations of $\frac{1}{2}$, $\frac{1}{4}$, and $\frac{1}{8}$ years. The tenor spacing in the model is $\boldsymbol{d} = 0.5$ with a flat initial forward curve of $F_k(0) = 6\%$ for all k. For the forward rate dynamics, we used one driving Brownian motion with a constant $\boldsymbol{l}_k(t) = 4.899\%$ for all k and t (roughly 20% in log-normal terms). The parameters for the stochastic volatility model were as follows: $V(0) = \boldsymbol{q} = 1$, $\boldsymbol{y}(x) = x^{3/4}$, $\boldsymbol{k} = 1$, and $\boldsymbol{e} = 1.4$. The dimensions of the ADI finite difference scheme used in the table were: 100 time steps, 150 steps in F, and 100 steps in V.

Appendix B
Details about the Expansion in Proposition III

We write the solution to the call option valuation problem as

$$G(t,F,V) = g\left(t,F;\overline{c}(t,V)\right) + \sum_{n=1}^{\infty} \frac{\partial^n g}{\partial c^n} \frac{E\left[\left(U(T) - \mu_U(t,V)\right)^n\right]}{n!\,t^n} \tag{B.1}$$

where $g(F,t;c)$ is given in Proposition I. The derivatives of g in (B.1) are evaluated at $c = \overline{c}(t,V)$, and the expectations are computed as

$$E\left[\left(U(T) - \mu_U(t)\right)^n\right] = (-1)^n \left.\frac{\partial^n l}{\partial s^n}\right|_{s=0};$$

see Lemma 2. From the result in the lemma we have, to order ε^4,

$$\begin{aligned}
E\left[\left(U(T) - \mu_U(t)\right)^2\right] &= 2\varepsilon^2 l_{1,2} + 2\varepsilon^4 l_{2,2} \\
E\left[\left(U(T) - \mu_U(t)\right)^3\right] &= -6\varepsilon^4 l_{2,3} \\
E\left[\left(U(T) - \mu_U(t)\right)^4\right] &= 24\varepsilon^4 l_{2,4}
\end{aligned}$$

so that (B.1) becomes

$$G(t,F,V) = g\left(t,F;\overline{c}(t,V)\right) + \tau^{-2}\left(\varepsilon^2 l_{1,2} + \varepsilon^4 l_{2,2}\right)\frac{\partial^2 g}{\partial c^2} - \varepsilon^4 \tau^{-3} l_{2,3}\frac{\partial^3 g}{\partial c^3} + \varepsilon^4 \tau^{-4} l_{2,4}\frac{\partial^4 g}{\partial c^4}, \tag{B.2}$$

to $O(\varepsilon^4)$. To compute (B.2) we need the derivatives of g with respect to c. From Proposition I we obtain, after some work,

$$\begin{aligned}
\frac{\partial g}{\partial c} &= F\phi(d_+)\Omega', \quad \frac{\partial^2 g}{\partial c^2} = R\frac{\partial g}{\partial c}, \\
\frac{\partial^3 g}{\partial c^3} &= R\frac{\partial^2 g}{\partial c^2} + R'\frac{\partial g}{\partial c}, \\
\frac{\partial^4 g}{\partial c^4} &= R\frac{\partial^3 g}{\partial c^3} + 2R'\frac{\partial^2 g}{\partial c^2} + R''\frac{\partial g}{\partial c},
\end{aligned} \tag{B.3}$$

where ϕ is the standard Gaussian density and primes denote derivatives with respect to c. Also,

$$
\begin{aligned}
R &= R(t,F,c) \equiv \Omega_{21} + d_+ d_- \Omega_{10}\,, \\
R' &= \Omega_{31} - \Omega_{21}^2 + d_+ d_- \Omega_{20} - \left(d_+^2 + d_-^2 + d_+ d_-\right)\Omega_{10}^2\,, \\
R'' &= \Omega_{41} - 3\Omega_{31}\Omega_{21} + 2\Omega_{21}^3 + d_+ d_- \Omega_{30} \\
&\quad -3\left(d_+^2 + d_-^2 + d_+ d_-\right)\Omega_{10}\Omega_{20} + 3\left(d_+ + d_-\right)^2 \Omega_{10}^3\,,
\end{aligned}
\tag{B.4}
$$

where $d_\pm$ is given in Proposition I, and

$$
\Omega_{mn} \equiv \frac{\partial^m \Omega / \partial c^m}{\partial^n \Omega / \partial c^n}\,.
$$

The derivatives of Ω with respect to c are listed below:

$$
\begin{aligned}
\partial\Omega/\partial c &= \tfrac{1}{2} c^{-1/2} t^{1/2} \Omega_0(F) + \tfrac{3}{2} c^{1/2} t^{3/2} \Omega_1(F), \\
\partial^2\Omega/\partial c^2 &= -\tfrac{1}{4} c^{-3/2} t^{1/2} \Omega_0(F) + \tfrac{3}{4} c^{-1/2} t^{3/2} \Omega_1(F), \\
\partial^3\Omega/\partial c^3 &= \tfrac{3}{8} c^{-5/2} t^{1/2} \Omega_0(F) - \tfrac{3}{8} c^{-3/2} t^{3/2} \Omega_1(F), \\
\partial^4\Omega/\partial c^4 &= -\tfrac{15}{16} c^{-7/2} t^{1/2} \Omega_0(F) + \tfrac{9}{16} c^{-5/2} t^{3/2} \Omega_1(F).
\end{aligned}
\tag{B.5}
$$

(B.2) represents a valid asymptotic expansion for $G(t,F,V)$. It is, however, often more convenient (and in many cases more accurate) to express the expansion in terms of implied "skew volatilities". For this, we write the option value as

$$
G(F,t,V) = g\left(F,t;\, c^*(t,V)\right) \tag{B.6}
$$

and expand $c^*(t,V)$ as

$$
c^*(t,V) = \bar{c}(t,V) + \varepsilon^2 c_1^*(t,V) + \varepsilon^4 c_2^*(t,V) + \ldots \tag{B.7}
$$

Substituting (B.7) into (B.6), expanding around $\bar{c}(t,V)$, and equating to (B.2) gives

$$
\begin{aligned}
c_1^*(t,V) &= t^{-2} l_{1,2} \bar{R}, \\
c_2^*(t,V) &= \frac{t^{-2} l_{2,2}\, \bar{g}' - t^{-3} l_{2,3}\, \bar{g}''' + \tfrac{1}{2} t^{-4} l_{1,2}^2 \left(\bar{g}^{(4)} - \bar{R}^2 \bar{g}'\right)}{\bar{g}'}\,,
\end{aligned}
\tag{B.8}
$$

where primes denote differentiation with respect to c, and overbars represent evaluation at $c = \bar{c}(t,V)$. We can now use (B.4) together with the result in Proposition I to obtain an expression for c_1^*, c_2^* as a polynomial in $Y = \ln(F/K)$. In particular, if we write

$$c_1^* = a_0 + a_1 Y^2,$$
$$c_2^* = b_0 + b_1 Y^2 + b_2 Y^4,$$

we get

$$a_0 = t^{-2} l_{1,2} \left(\Omega_{21} - \tfrac{1}{4} \Omega^2 \Omega_{10} \right),$$
$$a_1 = t^{-2} l_{1,2} \Omega^{-2} \Omega_{10},$$

and

$$b_0 = \begin{cases} t^{-2} l_{2,2} \left(\Omega_{21} - \tfrac{1}{4} \Omega^2 \Omega_{10} \right) - t^{-3} l_{2,3} \left(\Omega_{31} - \Omega_{21}^2 - \tfrac{1}{4} \Omega^2 \left(\Omega_{20} + \Omega_{10}^2 \right) + \left(\Omega_{21} - \tfrac{1}{4} \Omega^2 \Omega_{10} \right)^2 \right) \\ + \tfrac{1}{2} t^{-4} l_{1,2}^2 \left(\begin{array}{l} \Omega_{41} - 3\Omega_{31}\Omega_{21} + 2\Omega_{21}^3 - \tfrac{1}{4}\Omega^2 \Omega_{30} - \tfrac{3}{4}\Omega^2 \Omega_{10}\Omega_{20} \\ + 3\left(\Omega_{21} - \tfrac{1}{4}\Omega^2 \Omega_{10} \right) \left(\Omega_{31} - \Omega_{21}^2 - \tfrac{1}{4}\Omega^2 \left(\Omega_{20} + \Omega_{10}^2 \right) \right) \end{array} \right) \end{cases}$$

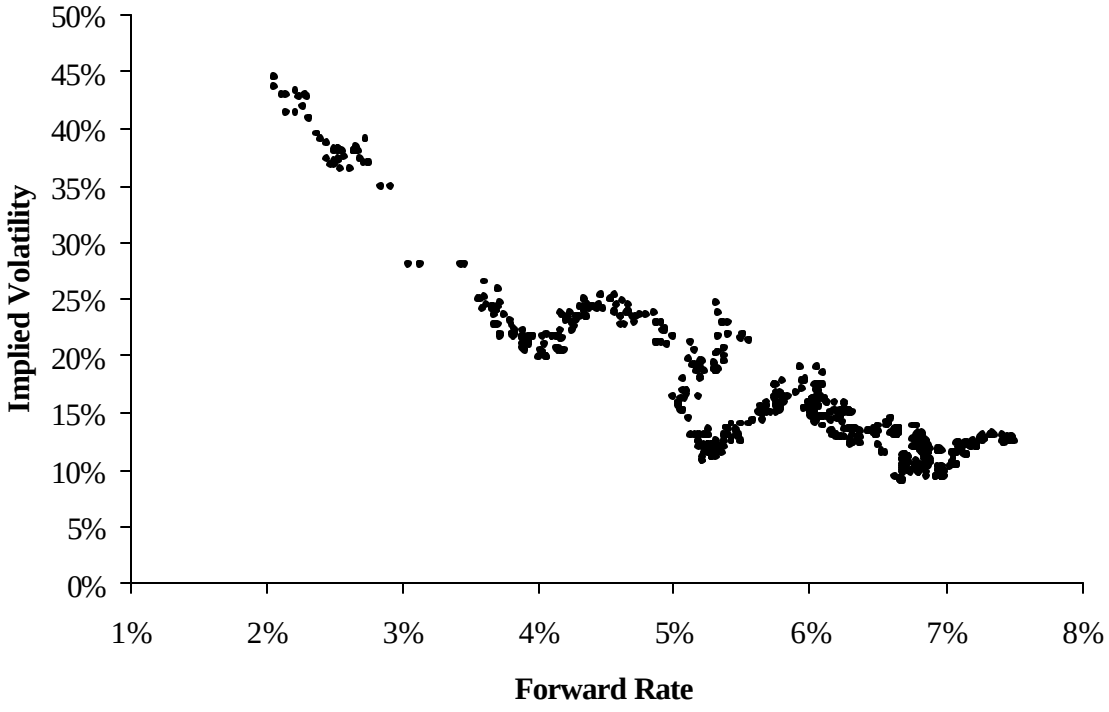
$$b_1 = \Omega^{-2} \begin{cases} t^{-2} l_{2,2} \Omega_{10} - t^{-3} l_{2,3} \left(\Omega_{20} - 3\Omega_{10}^2 + 2\Omega_{10} \left(\Omega_{21} - \tfrac{1}{4} \Omega^2 \Omega_{10} \right) \right) \\ + \tfrac{1}{2} t^{-4} l_{1,2}^2 \left(\begin{array}{l} \Omega_{30} - 9\Omega_{10}\Omega_{20} + 12\Omega_{10}^3 + 3\Omega_{10} \left(\Omega_{31} - \Omega_{21}^2 - \tfrac{1}{4}\Omega^2 \left(\Omega_{20} + \Omega_{10}^2 \right) \right) \\ + 3\left(\Omega_{21} - \tfrac{1}{4}\Omega^2 \Omega_{10} \right) \left(\Omega_{20} - 3\Omega_{10}^2 \right) \end{array} \right) \end{cases}$$

$$b_2 = \Omega^{-4} \left\{ -t^{-3} l_{2,3} \Omega_{10}^2 + \tfrac{3}{2} t^{-4} l_{1,2}^2 \Omega_{10} \left(\Omega_{20} - 3\Omega_{10}^2 \right) \right\}.$$

As the term $b_2 Y^4$ can be experimentally verified to cause a deterioration of the performance of the asymptotic expansion for large values of $|Y|$, we finally introduce the dampening factor $e^{-\Lambda e^2 Y^2}$, as listed in Proposition III. This extra step does not alter the order of the expansion.

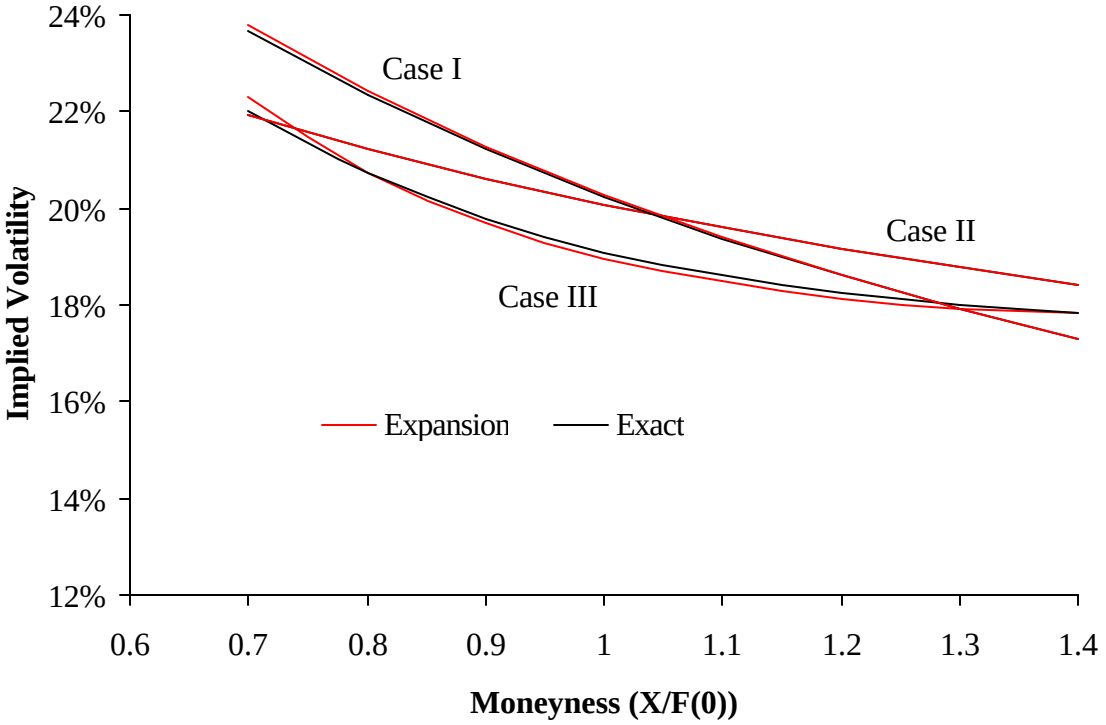
We round off this appendix by pointing out that that an expansion directly in terms of implied Black-Scholes volatility (rather than in the implied skew volatility) is also possible. Details are available upon request.

Figure 1: Implied 0.25-Year ATM Volatility of 1-Year US Forwards



Notes: The graph shows the observed implied (Black) volatility for 0.25-year US at-the-money option on 1-year rates, as a function of the forward rate itself. The time series in the graph is from Bloomberg and covers daily settings in the period 1/5/99 – 11/14/01.

Figure 2: Implied 10-year Caplet Volatility Skew in Selected DVF Models

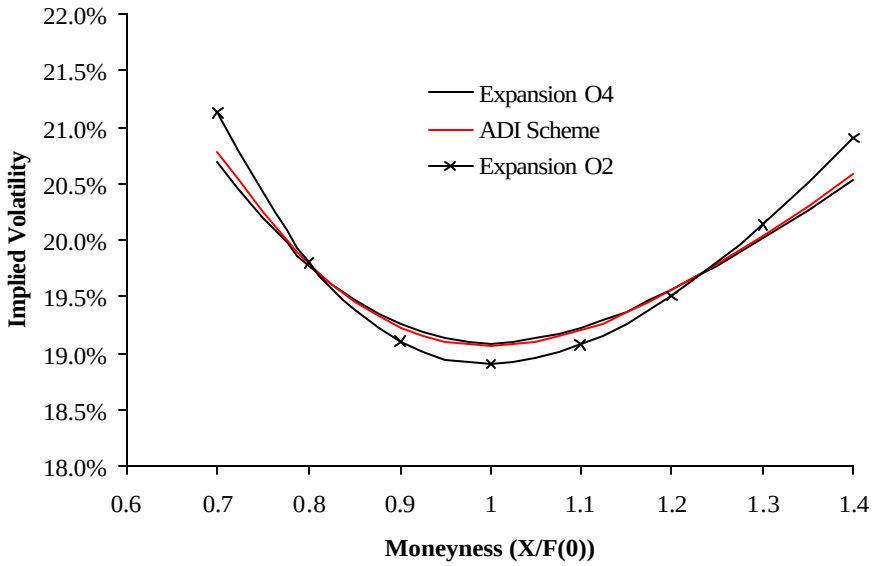


Notes: The graph shows the implied (Black) volatility for a 10-year caplet, as a function of the option moneyness factor $X / F(0)$. The stochastic volatility factor is assumed constant at $V = 1$, and $F(0) = 6\%$. The remaining model settings are:

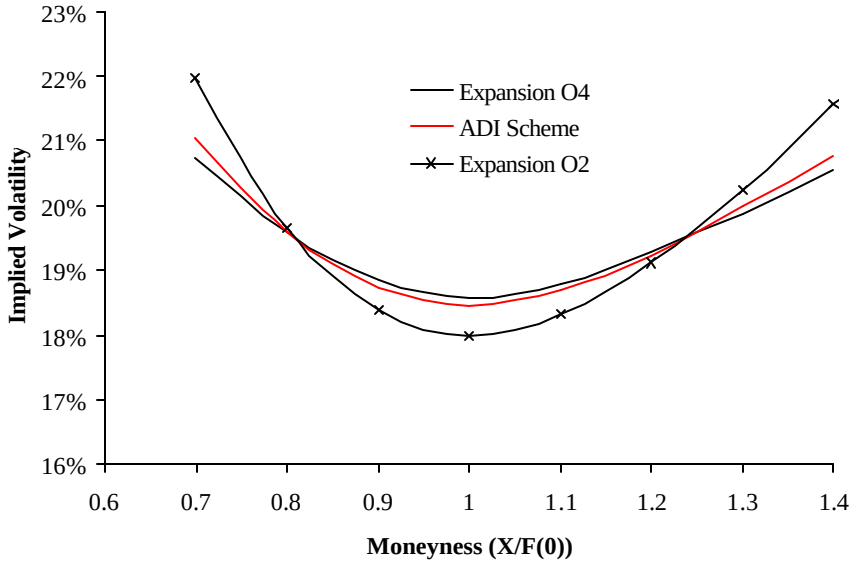
Case I: $j(x) = x^{0.1}$, $l = 1.59\%$.
Case II: $j(x) = \sqrt{x}$, $l = 4.90\%$.
Case III: $j(x) = x(1 + 30e^{-10x})$, $l = 16.75\%$.

The "Expansion" numbers in the graph are computed from the result in Proposition I. For Case I and Case II, the "Exact" numbers are computed by a known closed-form result (see e.g. Andersen and Andreasen (2000)); for Case III the "Exact" numbers are computed in a Crank-Nicholson finite difference grid with 150 time steps and 250 spatial steps. Note that for Case II, the "Exact" and "Expansion" numbers are indistinguishable to the resolution of the graph.

Figure 3: Implied 1.5-year Caplet Volatility Smile in Test A



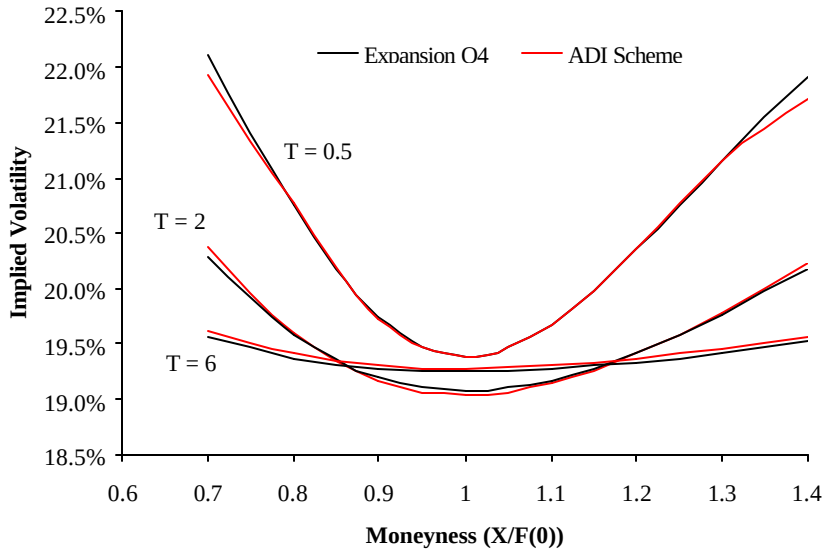
Panel A: $e = 150\%$, $k = 1$



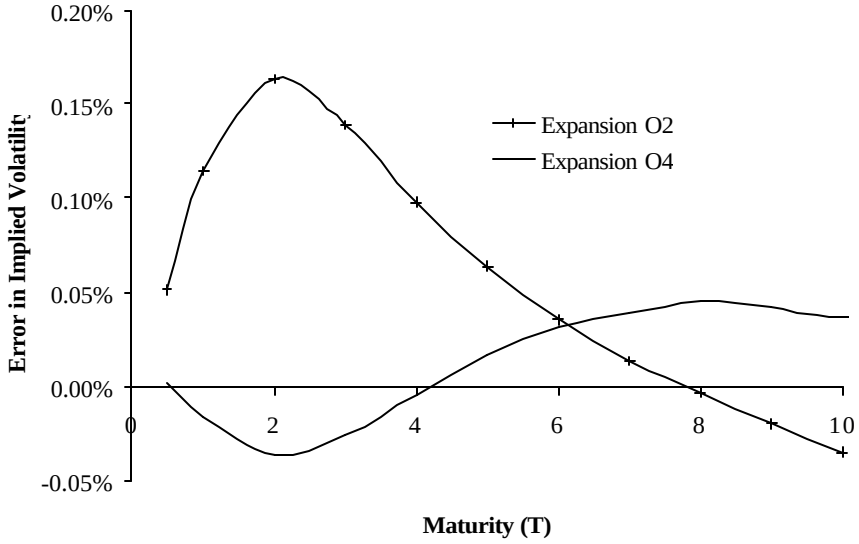
Panel B: $e = 200\%$, $k = 1$

Notes: Panels A and B show the implied (Black) volatility smile for a 1.5-year caplet in the Test A scenario with $k = 1$ and $e = 150\%$ and 200%, respectively. 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 with 100 time steps, 150 steps in F, and 100 steps in V.

Figure 4: Expansion Performance as a Function of Option Maturity in Test A



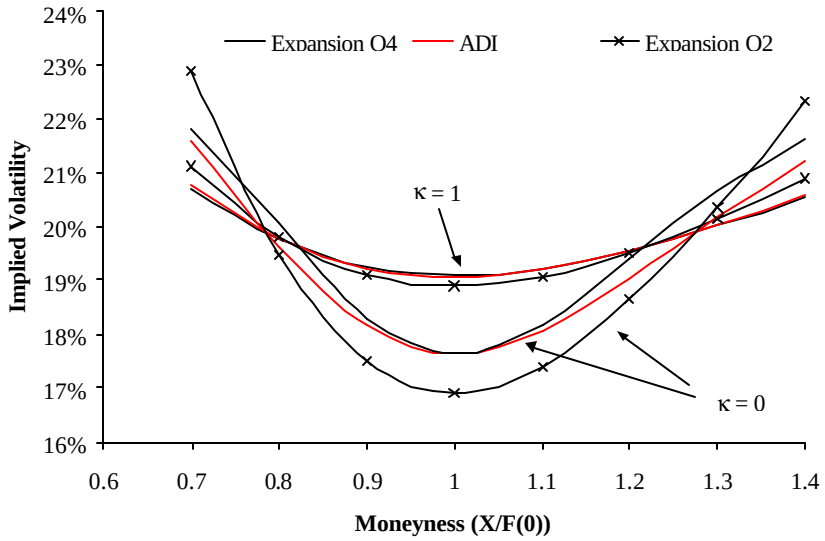
Panel A: Volatility Smiles



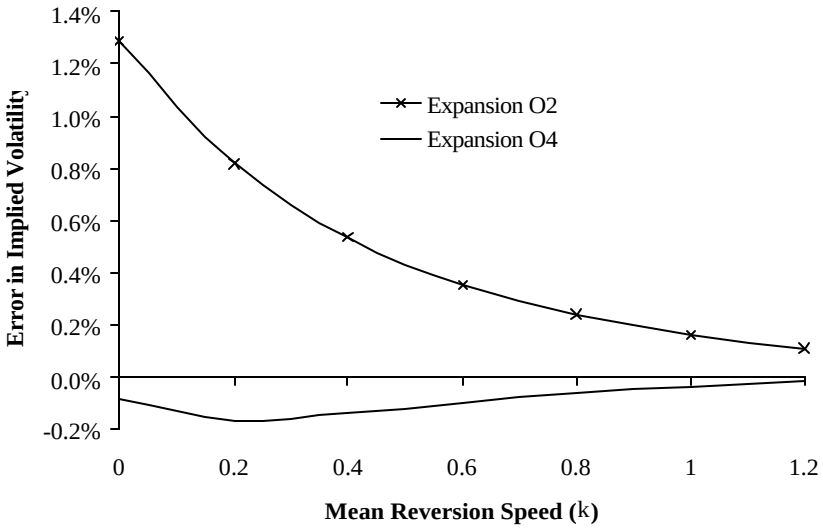
Panel B: Expansion Errors for At-the-Money Options

Notes: The graphs demonstrate the performance of the asymptotic expansions in Test A, with $e = 150\%$ and $k = 1$. Panel A shows volatility smiles at various maturities, as computed by an ADI scheme and expansion O4, respectively (expansion O2 was omitted for space considerations). Panel B displays the expansion errors (ADI implied volatility minus expansion implied volatility) for at-the-money options. In the computation of expansion O4 results, the cut-off parameter Λ in Eq. (21) was set to 1. The ADI numbers were computed with 150 time steps, 150 steps in F, and 100 steps in V.

Figure 5: Expansion Performance as a Function of Mean Reversion k in Test A



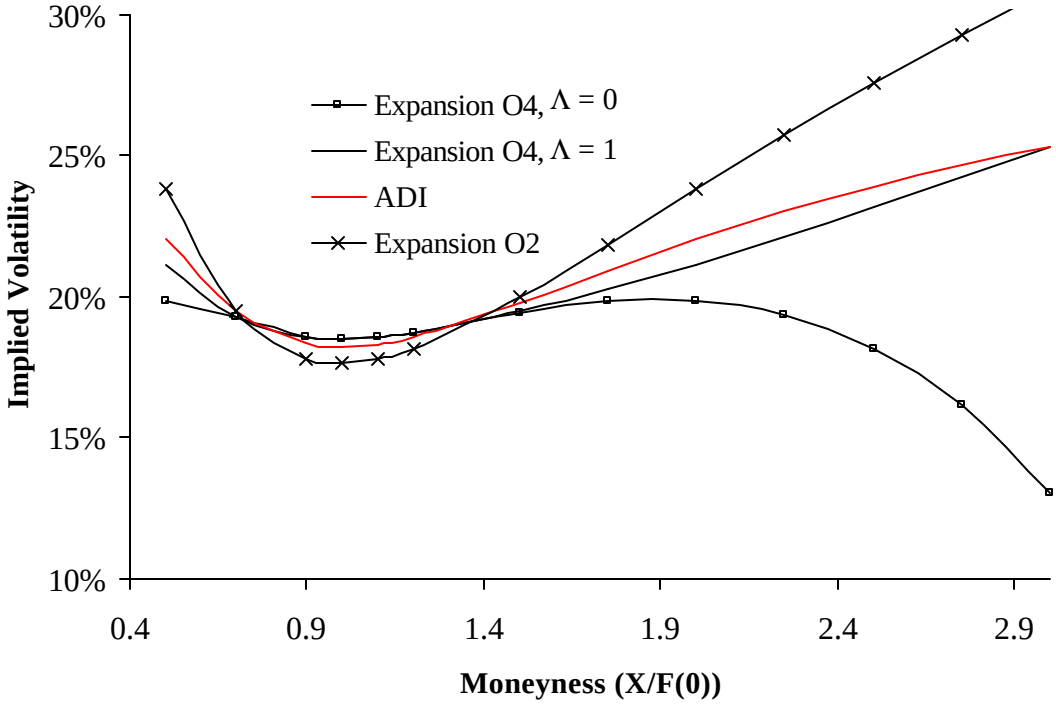
Panel A: Volatility Smiles; $T = 1.5$



Panel B: Expansion Error vs. k **;** $T = 2$

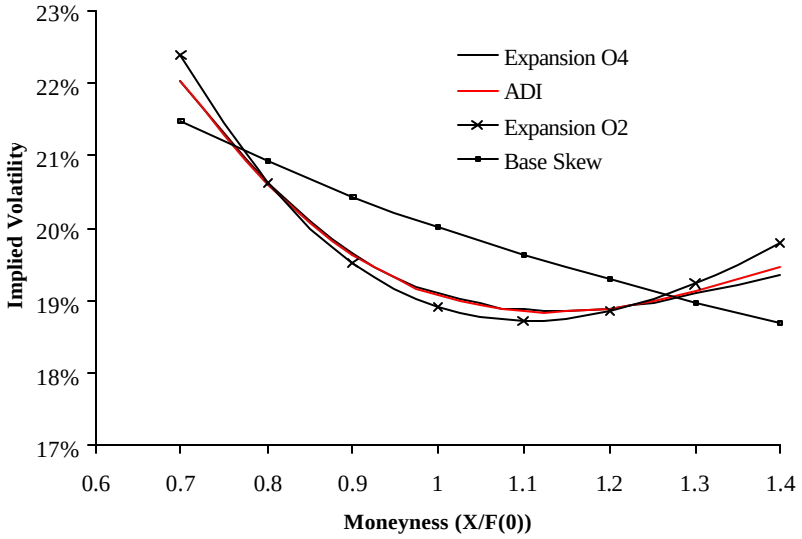
Notes: The graphs demonstrate the performance of the asymptotic expansions in Test A, with $e = 150\%$. Panel A shows the implied (Black) volatility smile for $T = 1.5$, and two different mean reversion speed levels : $k = 1$ and $k = 0$. Panel B displays the expansion errors (ADI implied volatility minus expansion implied volatility) for at-the-money options. In the computation of expansion O4 results, the cut-off parameter Λ in Eq. (21) was set to 1. The ADI numbers were computed with 150 time steps, 150 steps in F, and 100 steps in V.

Figure 6: Effect of Transcendental Dampening Factor Λ

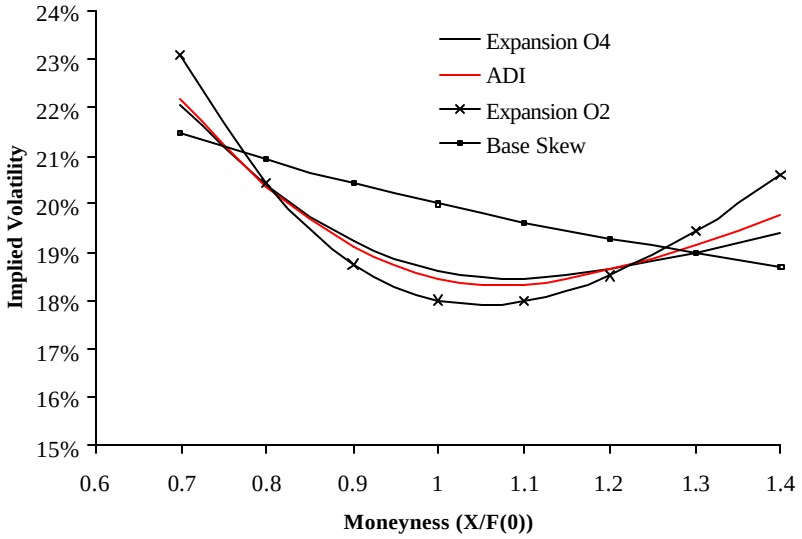


Notes: The graph shows the deeply in-the-money part of the implied (Black) caplet volatility smile for $T = 4$, $e = 150\%$, and $k = 0.5$. The remaining model parameters are as in the Test A scenario. The graph reports the results for expansion O4 using cut-off parameters of both $\Lambda = 1$ and $\Lambda = 0$. The ADI numbers reported in the graph were computed with 150 time steps, 200 steps in F, and 150 steps in V.

Figure 7: Implied 1.5-year Caplet Volatility Smile in Test B



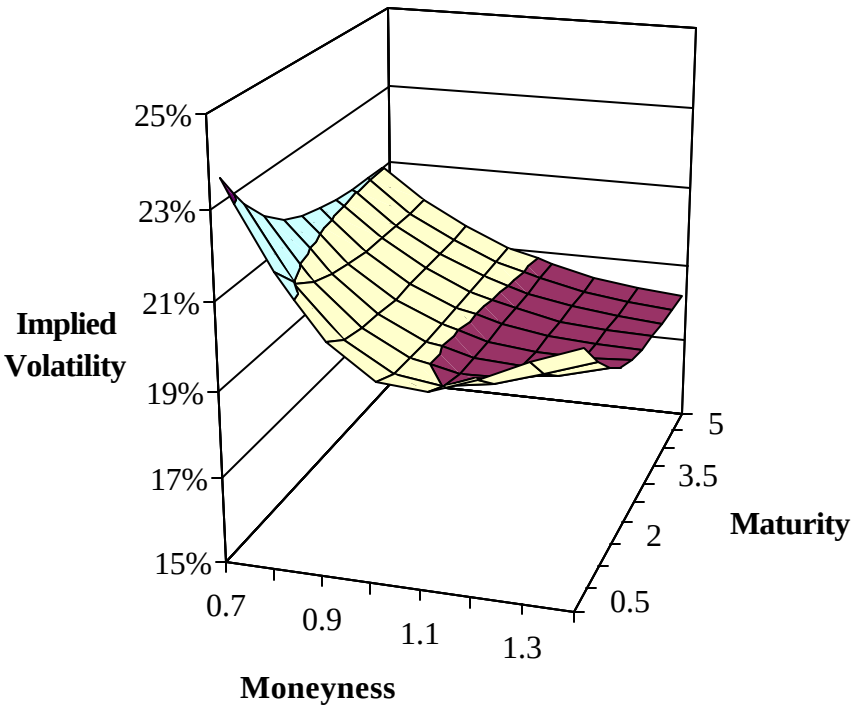
Panel A: $e = 150\%$, $k = 1$



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

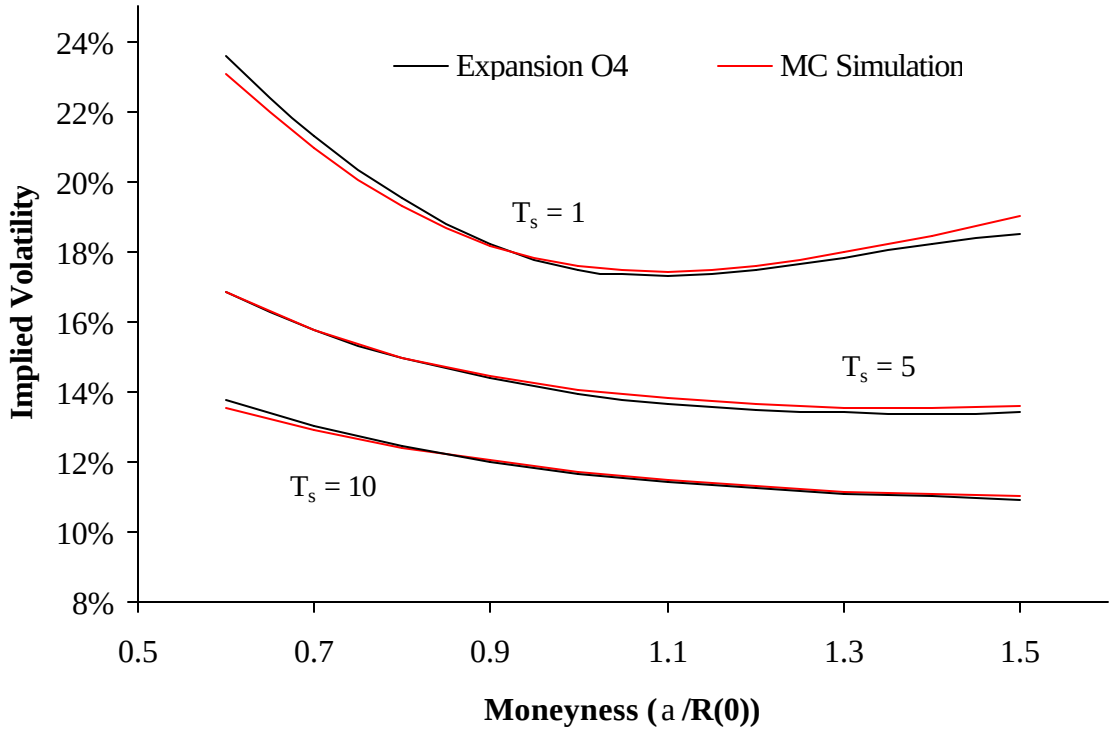
Notes: Panels A and B show the implied (Black) volatility smile for a 1.5-year caplet in the Test B scenario with $k = 1$ and $e = 150\%$ and 200%, respectively. In the computation of expansion O4 results, the cut-off parameter Λ in Eq. (21) was set to 1. The ADI numbers reported in the graph were computed with 100 time steps, 150 steps in F, and 100 steps in V. The "base skew" numbers correspond to $e = 0$ (no stochastic volatility component).

Figure 8: Volatility Smile Surface in Test B; $e = 150\%$, $k = 1$



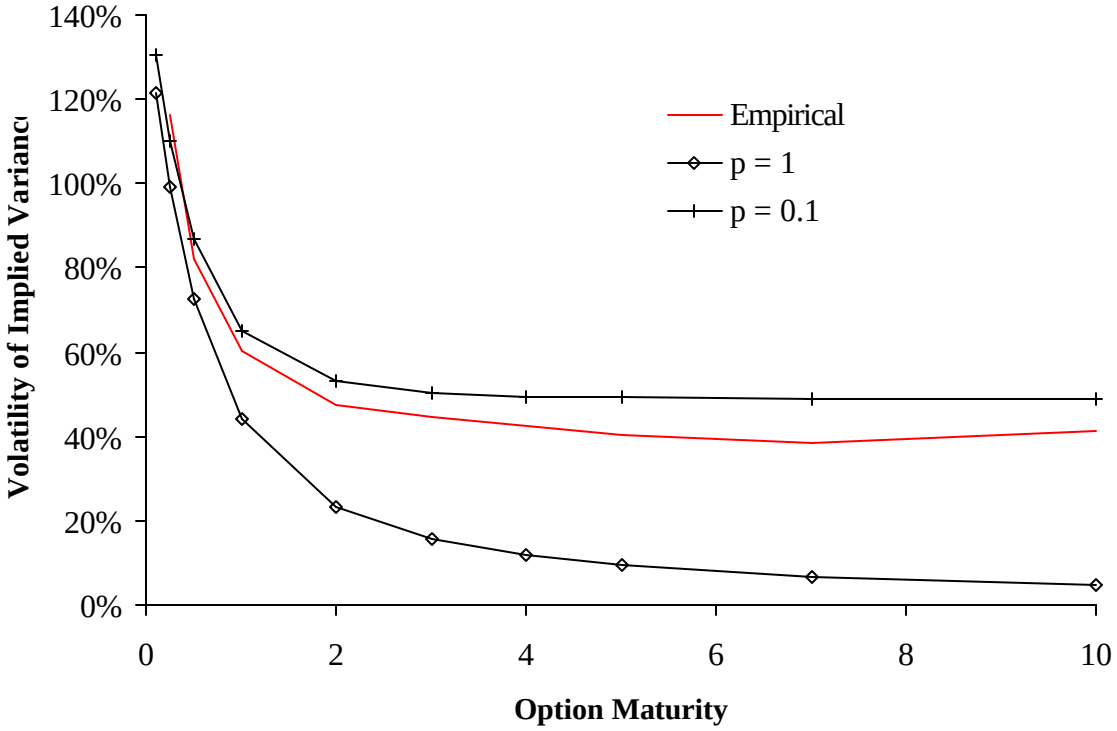
Notes: The graph shows the implied (Black) caplet volatility as a function of option maturity (T) and the option moneyness factor $X / F(0)$. $e = 150\%$, $k = 1$, and the remaining model parameters are as in the Test B scenario. All numbers in the graph were computed using Eqs. (21) and (22) (expansion O4) with the cut-off parameter Λ set to 1.

Figure 9: Implied Swaption Volatility Smile in Test C; 5-Year Swap Tenor



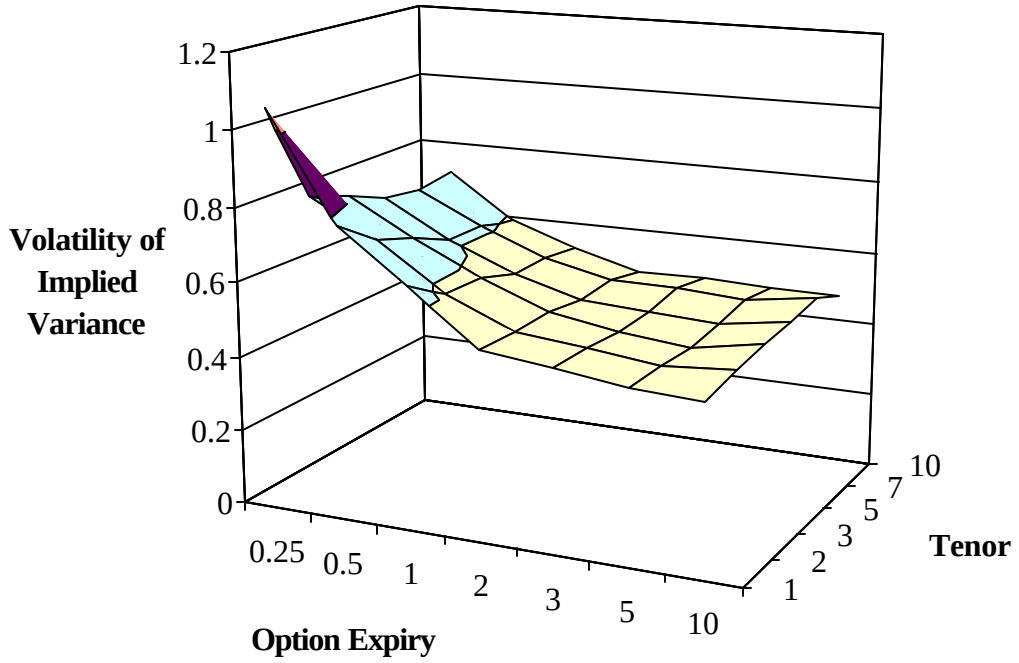
Notes: The graph shows the implied (Black) volatility smile for payer swaptions on a 5-year tenor swap (i.e. $T_e - T_s = 5$) for option maturities $T_s = 1$, $T_s = 5$, and $T_s = 10$. The moneyness factor in the graph is the ratio of the swap coupon to the par rate ($a / R(0)$). Model parameters are as in the Test C scenario. In the application of the expansion in Eq. (21), the cut-off parameter Λ was set to 1. The Monte Carlo results (MC Simulation) reported in the graph were computed with 50,000 antithetic paths (a total of 100,000 separate paths) and the simulation algorithm (A.1)-(A.2) in Appendix A with a constant time step of $\Delta = \frac{1}{8}$. The average sample error (in terms of implied Black volatility) was around 0.1%.

Figure 10: Volatility of Implied Variance of 1-Year Rate in US



Notes: The graphs shows the volatility of implied variance for US swaptions on a 1-year rate, as estimated empirically and as computed in two different models. The empirical estimation is based on two years of weekly data from Bloomberg. The model data uses $e = 130\%$, $k = 3$, $y(x) = x^{3/4}$, $F(0) = 6\%$, and $lj\left(F(0)\right) = 1.2\%$. The skew function is $j(x) = x^p$, with p being either 0.1 or 1.

Figure 11: Volatility of Implied Variance of Various Rates in US



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