Stochastic Volatility Corrections for Interest Rate Derivatives

Peter Cotton[*]   Jean-Pierre Fouque[†]   George Papanicolaou[‡]

Ronnie Sircar[§]

First Draft: May 2000; Final Version: February 2003

Abstract

We study simple models of short rates such as the Vasicek or CIR models, and compute corrections that come from the presence of fast mean-reverting stochastic volatility. We show how these small corrections can affect the shape of the term structure of interest rates giving a simple and efficient calibration tool. This is used to price other derivatives such as bond options. The analysis extends the asymptotic method developed for equity derivatives in (Fouque, Papanicolaou and Sircar 2000b). The assumptions and effectiveness of the theory are tested on yield curve data.

Contents

1 Introduction

In this article we model and exploit the bursty nature of interest-rate volatility to try to understand, in a relatively robust manner, the *effect* of uncertain and changing volatility on rate-sensitive derivative instruments. A key goal is to move away from the affine family of models and capture a wider range of observed yield curve shapes, while staying within a parsimonious class of stochastic volatility models.

Our empirical evidence, described in Section 5, suggests that interest rates and interest rate volatility vary on different time scales, the latter exhibiting a characteristic mean-reversion time on the order of two weeks. We are able to capture stochastic volatility effects with minimal calibration overhead. The ability to fit unusual yield curves during the bond crisis of 1998 is striking.

We begin with a review of the motivation for stochastic volatility models for the short-rate, and describe the observation of volatility persistence in terms of rates of mean-reversion. This sets up

the analysis by which contingent claims prices are approximated by an asymptotic series of terms of decreasing magnitude. Truncation of the series at a judicious point reduces the dependence of the approximation on the details of the underlying model. In this sense, the approximation is *robust*, being valid for a large class of models. Introducing more terms in the series would increase the dependence on a specific model, while a fewer number of terms misses important features, like humps in the yield curve, seen in the data. This approach was introduced to study the effects of implied volatility skews on equity derivatives in (Fouque et al. 2000b).

The analysis identifies the important groupings of the base model parameters that summarize, to the maintained level of accuracy, the market's view of volatility risk. We find from empirical analysis of short rate data a "bursty" volatility with a natural time scale on the order of a couple of weeks.

The short rate itself may be considered observable if we allow the one-month rate to serve as a proxy, for example. The present volatility level is unobservable, but to the level of approximation we shall use, it does not enter the prices. The analysis reduces dependence on the hidden factors while maintaining a reasonably rich dynamic for the short rate. The approach described here lends itself to robust estimation techniques which are less computationally intensive than likelihood or moment methods, facilitating studies of parameter stability, model risk, hedging and risk analysis.

2   Background

We proceed here in the tradition of modeling at the level of the short-rate process $(r_t)_{t \geq 0}$ from which the yield curve and derivatives prices will be determined by the principle of *no arbitrage*. One-factor models, in which (r_t) is an Itô process driven by a single Brownian motion date back at least to (Vasicek 1977). However, these models imply perfect correlation between returns on bonds of all maturities, which is contrary to empirical observation and, as a result, there has been much work on multifactor models. We refer to surveys of the literature in (Hull 1999, Chapter 21) or (Duffie 2001, Chapter 7).

2.1   Motivation for Stochastic Volatility Models

Stochastic volatility models, in which a second random process drives the diffusion term of the short-rate, are popular examples of two-factor models. The second factor may not actually be the instantaneous standard deviation of short-rate rate changes, but we will loosely refer to it as the volatility-driving process. As explained in (Longstaff and Schwartz 1992), there is much empirical evidence of randomly changing volatility, and, as with equities, volatility is an important variable in determining derivative prices. In fact, stochastic volatility models have been successfully used to explain the observed implied volatility surface or skew in equity markets. See (Frey 1996, Fouque et al. 2000b) for example. Models of interest rate stochastic volatility are often motivated by their *a posteriori* impact on explaining term-structure shapes and movements (Longstaff and Schwartz 1992).

For an overview of issues of model selection, we refer to (Rogers 1995) and references therein.

2.2   Calibration & Separation of Scales

The computational cost of introducing a second factor is high in practical situations and this makes it preferable to choose models for which vanilla (liquid) instruments have closed-form formulas. Parameters are then calibrated or "backed-out" from, say, the observed yield curve. A popular

model, which includes stochastic volatility effects, is the generalized two-factor CIR, studied in (Longstaff and Schwartz 1992) and (Duffie and Kan 1996), which we write

$$dr_t = \kappa_1(\mu_1 - r_t)dt + \sqrt{c_{11}r_t + c_{12}V_t}\, dW_t^\star, \tag{1}$$

$$dV_t = \kappa_2(\mu_2 - V_t)dt + \sqrt{c_{21}r_t + c_{22}V_t}\, dZ_t^\star. \tag{2}$$

Here $(W_t^\star, Z_t^\star)$ are Brownian motions under the pricing measure, and in fact the drift terms can be linear in both processes (r_t) and (V_t) in the most general form of the model.

The advantages of this model are that $(V_t)_{t\geq 0}$ can be interpreted as a volatility-driving process that is mean-reverting, the solution (r_t) stays positive (with suitably restricted parameter values) and there is an affine formula for the yield curve, in which the time-dependent coefficients are solutions of ODEs, simplifying the inverse problem of calibration.

A different way to introduce volatility randomness, mean-reversion and clustering, is to take the volatility to be some positive function $f(Y_t)$ of a mean-reverting factor process $(Y_t)_{t\geq 0}$. A generic example of such a process is the Ornstein-Uhlenbeck (OU) process, satisfying the SDE

$$dY_t = \alpha(m - Y_t)dt + \beta d\hat{Z}_t,$$

with $(\hat{Z}_t)$ a Brownian motion. In this family of models, (Y_t) is an autonomous diffusion. What is important for the asymptotic results we present is that (Y_t) is an ergodic process with a unique invariant distribution, modeling volatility mean-reversion. In the OU case the invariant distribution is normal $\mathcal{N}(m, \nu^2)$, where we define $\nu^2 = \beta^2/2\alpha$.

The generality of the class of models is in not having to specify a function f: the features of this function that are needed for the theory are captured by *group parameters* derived in the asymptotics and easily calibrated from data.

The asymptotic approximations we present are in the limit $\alpha \to \infty$ with ν^2 fixed, which we refer to as *fast mean-reversion*. This method has been used to robustly incorporate stochastic volatility effects into equity derivative prices to explain the implied volatility skew (Fouque et al. 2000b), and from the skew price other contracts such as barriers (Fouque, Papanicolaou and Sircar 1999) or American options (Fouque, Papanicolaou and Sircar 2001). The group market parameters identified by the analysis have been found to be relatively stable over time when estimated from market data (Fouque, Papanicolaou and Sircar 2000a).

3   Vasicek Model: Notation & Brief Review

In this section, we describe the simple one-factor Vasicek model (Vasicek 1977) and review how bonds and bond options are priced under it. Details can be found in (Jamshidian 1989). Stochastic volatility is introduced into this model in the next section. The choice of this "background" model is motivated by simplicity to illustrate the method. The analogous calculation built around a CIR model is given in Appendix A.

3.1   Model

In the Vasicek model, the short-rate is modeled as a mean-reverting Gaussian stochastic process $(\bar{r}_t)_{t\geq 0}$ on a probability space $(\Omega, \mathcal{F}, I\!\!P)$ equipped with a filtration $(\mathcal{F}_t)_{t\geq 0}$. Under the real-world probability measure $I\!\!P$, it satisfies the linear stochastic differential equation (SDE)

$$d\bar{r}_t = a(\bar{r}_\infty - \bar{r}_t)dt + \bar{\sigma}d\bar{W}_t, \tag{3}$$

where $(\bar{W}_t)_{t \geq 0}$ is a standard Brownian motion. Here $\bar{\sigma}$ is a constant volatility, assumed positive, and a is the rate of mean-reversion, also assumed positive.

We use the overbar notation to highlight the one-factor world, to which we shall make reference in the next section as well as when we give the asymptotic approximations for the two-factor model there.

Under an equivalent martingale (pricing) measure $I\!\!P^{\star}$, it also follows a linear SDE

$$d\bar{r}_t = a(r^{\star} - \bar{r}_t)dt + \bar{\sigma}d\bar{W}_t^{\star}\,, \tag{4}$$

where $(\bar{W}_t^{\star})$ is a standard $I\!\!P^{\star}$-Brownian motion, if we assume the market price of interest rate risk, denoted by λ, to be constant. It is included in $r^{\star} = \bar{r}_{\infty} - \frac{\lambda\bar{\sigma}}{a}$. In other words, in the risk-neutral world $I\!\!P^{\star}$, the short rate $(\bar{r}_t)$ is an Ornstein-Uhlenbeck process fluctuating around its mean level $r^{\star}$ with a rate of mean-reversion a, assumed positive. It has the undesirable feature that it can become negative, but it is the simplest model in a larger family of models known as *affine* for which the computation of bond prices and derivatives is relatively easy.

3.2   Bond Prices

The no-arbitrage price at time t of a zero-coupon bond maturing at time T, denoted by $L(t, T)$, is given by

$$L(t, T) = I\!\!E^{\star}\left\{e^{-\int_t^T \bar{r}_s\,ds} \mid \mathcal{F}_t\right\} = \bar{P}(t, \bar{r}_t; T), \tag{5}$$

where the bond pricing function $\bar{P}(t, x; T)$ satisfies the partial differential equation (PDE)

$$\frac{\partial \bar{P}}{\partial t} + \frac{1}{2}\bar{\sigma}^2\frac{\partial^2 \bar{P}}{\partial x^2} + a(r^{\star} - x)\frac{\partial \bar{P}}{\partial x} - x\bar{P} = 0\,, \tag{6}$$

with the terminal condition $\bar{P}(t, x; T) = 1$.

The solution is

$$\bar{P}(t, x; T) = A(T - t)e^{-B(T-t)x}\,, \tag{7}$$

where A and B satisfy the system of ordinary differential equations

$$B' + aB - 1 = 0 \tag{8}$$

$$-\frac{A'}{A} + \frac{1}{2}\bar{\sigma}^2 B^2 - ar^{\star}B = 0, \tag{9}$$

with $A(0) = 1$ and $B(0) = 0$. This gives explicitly

$$B(\tau) = \frac{1 - e^{-a\tau}}{a} \tag{10}$$

$$A(\tau) = \exp\left\{-\left[R_{\infty}\tau - R_{\infty}\frac{1 - e^{-a\tau}}{a} + \frac{\bar{\sigma}^2}{4a^3}\left(1 - e^{-a\tau}\right)^2\right]\right\}\,, \tag{11}$$

where we have set

$$R_{\infty} = r^{\star} - \frac{\bar{\sigma}^2}{2a^2} = \bar{r}_{\infty} - \frac{\lambda\bar{\sigma}}{a} - \frac{\bar{\sigma}^2}{2a^2}\,.$$

The *yield curve*, defined by $R(t, \tau) = -\frac{1}{\tau}\log\left(L(t, t + \tau)\right)$, as a function of τ, is

Figure 1 shows a typical bond pricing function (of maturity) and the corresponding yield curve obtained under a Vasicek model. The yield curve (bottom) is a classical increasing curve obtained with this model when the current short rate is lower than the long term rate.
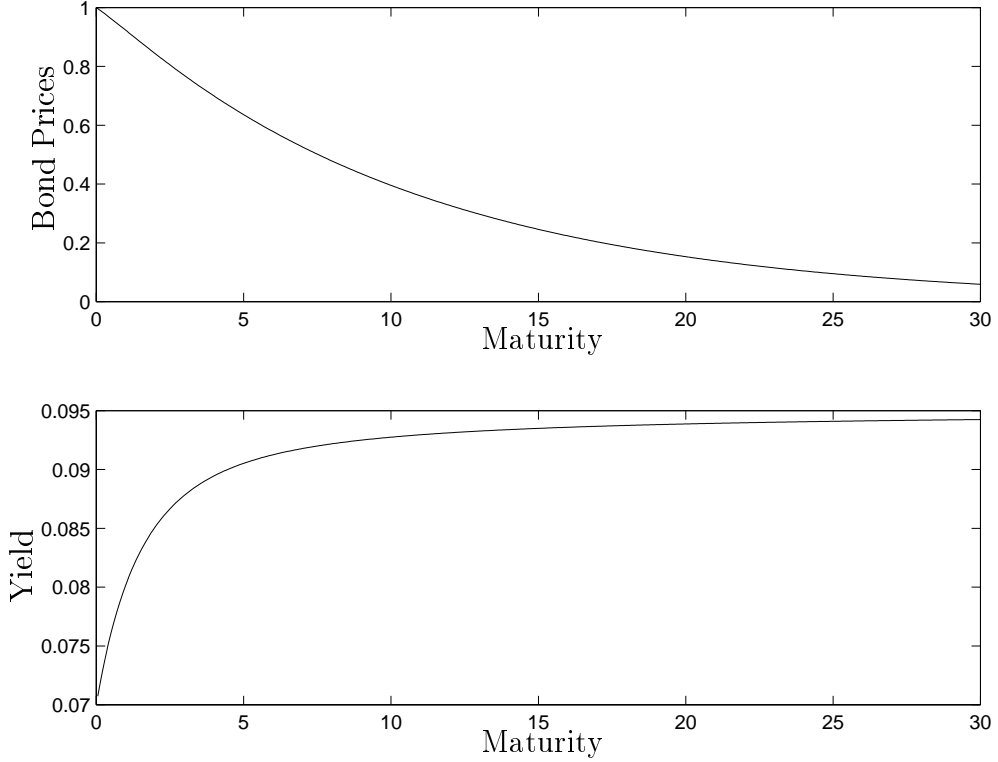
5

Figure 1: *Bond prices (top) and Yield curve (bottom) in the Vasicek model with $a = 1$, $r^\star = 0.1$ and $\bar{\sigma} = 0.1$. Maturity runs from 0 to 30 years. $R_\infty = 0.095$ and the initial rate is $x = 0.07$.*

3.3 Bond Option Prices

We continue to denote the maturity of the bond by T and the maturity of a European option written on that bond is denoted by T_0 with $t \leq T_0 < T$. The payoff of the option at time T_0 is a function $h(L(T_0, T))$ of the bond price at the expiration time of the option.

The no-arbitrage price of this option is given by $\bar{Q}(t, x; T, T_0)$, solution of the partial differential equation (6) with the terminal condition $\bar{Q}(T_0, x; T, T_0) = \tilde{h}(x) = h\left(\bar{P}(T_0, x; T)\right)$ where $\bar{P}(T_0, x; T)$ is given by (7) at $t = T_0$.

In particular, the price $\bar{C}(t, x)$ of a call option with strike price K, is given by

$$\bar{C}(t, x) = \bar{P}(t, x; T)N(h_1) - K\bar{P}(t, x; T_0)N(h_2), \tag{12}$$

where N is the $\mathcal{N}(0, 1)$ cumulative distribution function. The quantities $h_{1,2}$ are given explicitly by

$$h_{1,2} = \frac{\log\left(\frac{\bar{P}(t,x;T)}{\bar{P}(t,x;T_0)}\right) - \log K \pm \frac{1}{2}v^2}{v}, \tag{13}$$

$$v^2 = \frac{\bar{\sigma}^2}{2a^3}\left(1 - e^{-2a(T_0-t)}\right)\left(1 - e^{-a(T-t)}\right)^2. \tag{14}$$

4 Stochastic Volatility Vasicek Models

In this section, we model stochastic volatility in the short-rate and present the asymptotic analysis for bond and bond option prices that corrects the Vasicek formulas of the previous section to

6

account for uncertain volatility. The key assumption is that volatility is fast mean-reverting, as defined in Section 2.2.

4.1 Model

We extend model (4) by introducing stochastic volatility σ_t given by a positive function $f(Y_t)$ of an ergodic Markov diffusion process (Y_t) with a unique invariant distribution, for example an Ornstein-Uhlenbeck process. The Gaussian distribution of this last choice is not important for the asymptotic results as we shall calibrate them, but we write a specific driving process to fix ideas.
Assumption The volatility function f is smooth and bounded above and below:

$$0 < c_1 \le f \le c_2 < \infty \tag{15}$$

for some constants c_1 and c_2, and f has bounded derivatives.

Under the real world measure $I\!\!P$ the processes (r_t, Y_t) satisfy

$$
\begin{aligned}
dr_t &= a(r_\infty - r_t)\, dt + f(Y_t)\, dW_t \\
dY_t &= \alpha(m - Y_t)\, dt + \beta\left(\rho\, dW_t + \rho'\, dZ_t\right),
\end{aligned}
\tag{16}
$$

where (W_t) and (Z_t) are independent standard Brownian motions, and $\rho' = \sqrt{1 - \rho^2}$. The parameter ρ with $|\rho| < 1$ allows a correlation between the Brownian motion (W_t) driving the short rate and its volatility. Typically, we would expect $\rho > 0$ as rising volatility tends to push bond prices down and yields up. This is confirmed from empirical evidence. Observe that the invariant distribution of (Y_t) is the normal distribution $\mathcal{N}(m, \beta^2/2\alpha)$.

In the risk-neutral world $I\!\!P^{\star(\lambda,\gamma)}$, the model for the short rate (r_t) becomes

$$
\begin{aligned}
dr_t &= \left(a(r_\infty - r_t) - \lambda(Y_t)f(Y_t)\right)\, dt + f(Y_t)\, dW_t^\star \tag{17} \\
dY_t &= \left(\alpha(m - Y_t) - \beta\left[\rho\lambda(Y_t) + \rho'\gamma(Y_t)\right]\right)\, dt + \beta\left(\rho dW_t^\star + \rho'\, dZ_t^\star\right), \tag{18}
\end{aligned}
$$

where $(W_t^\star)$ and $(Z_t^\star)$ are two independent standard $I\!\!P^{\star(\lambda,\gamma)}$-Brownian motions.

The market price of risk $\lambda(Y_t)$ may depend on Y_t but we assume that it does not depend on the short rate r_t. Similarly the *market price of volatility risk* $\gamma(Y_t)$ associated with the second source of randomness (Z_t) may depend on Y_t or simply be a constant. The effect is to preserve under the pricing measure the Markovian structure of the model and the autonomy of the volatility-driving process (Y_t), so that (17-18) remains a genuine stochastic volatility model. The assumption is that the market does not price term-structure derivatives as though the volatility is significantly affected by the level of the short-rate, or non-Markovian phenomena.
Assumption The risk premium processes $\lambda(Y_t)$ and $\gamma(Y_t)$ are bounded:

$$c_1' \le \lambda, \gamma \le c_2' < \infty \tag{19}$$

for some constants c_1' and c_2', and λ and γ have bounded derivatives.

For general functions f these models do not belong to the affine family in that yields will not be affine in the second factor Y_t. In fact they do remain affine in the short-rate r_t, which we shall exploit in the following sections. However explicit formulas for bond prices through ordinary differential equations are not available.

4.2 Correction for Bond Prices

The no-arbitrage price $L(t, T)$ of a zero-coupon bond maturing at time T is now given by

$$P(t, x, y; T) = I\!\!E^{\star(\lambda, \gamma)} \left\{ e^{-\int_t^T r_s ds} | r_t = x, Y_t = y \right\}, \tag{20}$$

where the expectation $I\!\!E^{\star(\lambda, \gamma)}$ is taken with respect to the distribution of (r_t, Y_t) solution of (17-18) starting at time t from (x, y). Note that in these models, (r_t) is not bounded below so the expectation in (20) is not trivially finite. In the next section, we prove that in fact it *is* finite and that P is a classical solution of the Feynman-Kac partial differential equation

$$\frac{\partial P}{\partial t} + \frac{1}{2} f(y)^2 \frac{\partial^2 P}{\partial x^2} + (a(r_\infty - x) - \lambda(y) f(y)) \frac{\partial P}{\partial x} - xP$$
$$+ \beta \rho f(y) \frac{\partial^2 P}{\partial x \partial y} + \frac{1}{2} \beta^2 \frac{\partial^2 P}{\partial y^2} + (\alpha(m - y) - \beta \Lambda(y)) \frac{\partial P}{\partial y} = 0, \tag{21}$$

with the terminal condition $P(T, x, y; T) = 1$ for every x and y, where we define

$$\Lambda(y) = \rho \lambda(y) + \rho' \gamma(y),$$

the combined market price of risk.

4.2.1 Properties of the Bond Pricing Function

First, we define

$$M(t, y) = I\!\!E \left\{ e^{\int_t^T c(s, \tilde{Y}_s)} \mid \tilde{Y}_t = y \right\},$$

where $\tilde{Y}$ is an Itô process defined by

$$d\tilde{Y}_t = \left(\alpha(m - \tilde{Y}_t) - \beta b(t, \tilde{Y}_t) \right) dt + \beta \, d\tilde{W}_t,$$

on some probability space where $\tilde{W}$ is a standard Brownian motion and we define

$$b(t, y) = \Lambda(y) + \rho f(y) B(T - t) \tag{22}$$
$$c(t, y) = \frac{1}{2} f(y)^2 B(T - t)^2 - B(T - t)(ar_\infty - \lambda(y) f(y)). \tag{23}$$

Under our assumptions (15) and (19) on $f(y)$, $\lambda(y)$ and $\gamma(y)$, and the form of B given by (10), the coefficients of $\tilde{Y}$ are smooth with bounded derivatives. The potential $c(t, y)$ is bounded in y and $t \leq T$, and therefore $M(t, y)$ is also bounded. Moreover, M is the unique classical solution of the Feynman-Kac partial differential equation:

$$\frac{\partial M}{\partial t} + \frac{1}{2} \beta^2 \frac{\partial^2 M}{\partial y^2} + (\alpha(m - y) - \beta b(t, y)) \frac{\partial M}{\partial y} + c(t, y) M = 0, \tag{24}$$
$$M(T, y) = 1. \tag{25}$$

This follows from (Krylov 1980, Theorem 2.9.10). See also (Duffie 2001, Appendix E). In these references, it is assumed that $c \leq 0$, but in our case, c is bounded above by some $c_0 < \infty$, so we apply the theorem to $e^{-c_0(T-t)} M$ which satisfies the same PDE with c replaced by $c - c_0$, which fulfills the required condition on the potential term.

Then we define P_M by

$$P_M(t, x, y; T) = M(t, y) e^{-B(T-t)x}, \tag{26}$$

where B is given by (10). It follows from direct calculation that P_M is a classical solution of (21) with terminal condition 1. The following proposition shows that $P_M = P$.

Proposition 1 *The bond price (20) is given by formula (26). It is exponential-affine in the x-variable:*

$$\mathbb{E}^{\star(\lambda,\gamma)} \left\{ e^{-\int_t^T r_s \, ds} \mid r_t = x, Y_t = y \right\} = M(t,y)e^{-B(T-t)x}. \tag{27}$$

Proof. From (17), and the boundedness assumptions on λ and f, (r_t) satisfies an equation of the form

$$dr_t = -ar_t \, dt + g_t \, dt + k_t \, dW_t^{\star},$$

where $|g| \leq C_1$, and $|k| \leq C_2$ are bounded ($\mathcal{F}_t$-adapted) processes. Then, for any $T > 0$, we have that $\sup_{0 \leq t \leq T} |r_t|$ is stochastically dominated by

$$|r_0| + \frac{C_1}{a} + \left| \tilde{W}_{\left(C_2^2 e^{2aT}/a\right)} \right|, \tag{28}$$

where $\tilde{W}$ is a standard Brownian motion. Therefore r_t and $\int_0^t r_s \, ds$ have exponential moments (uniform in $\varepsilon > 0$), and the finiteness of the bond price given by the expected value (20) follows. Using the form (26) of the solution of the Feynman-Kac PDE (21) we deduce that

$$e^{-\int_0^t r_s \, ds} P_M(t, r_t, Y_t; T) = M(t, Y_t)e^{-B(T-t)r_t} e^{-\int_0^t r_s \, ds}$$

is a martingale. Integrability is given by the exponential moments above. Itô's formula and the PDE (24) satisfied by $M(t,y)$ give the conclusion. When applying Itô's formula, e^{-Bx} is not bounded in x but it is justified by the exponential moments. The martingale property between t and T and the Markov property of the pair (r_t, Y_t) give the formula (27). Hence $P_M = P$.

Observe that in the class of stochastic volatility models we consider, the yield is affine in the short rate r_t but is general in the other factor Y_t.

4.2.2   Scales in Stochastic Volatility

In the context of fast mean-reverting stochastic volatility, the rate of mean-reversion α in the OU process (Y_t) driving the volatility is large and we are interested in approximations in the limit $\alpha \to \infty$ with $\nu = \beta/\sqrt{2\alpha}$ remaining constant.

We use the notation and the method we have developed in the context of equity markets in (Fouque et al. 2000b). The asymptotic analysis will be carried out as $1/\alpha$ becomes small and we set:

$$\begin{aligned} \varepsilon &= 1/\alpha, \\ \beta &= \frac{\nu\sqrt{2}}{\sqrt{\varepsilon}}, \\ P(t,x,y;T) &= M^\varepsilon(T-t,y)e^{-B(T-t)x}, \tag{29} \end{aligned}$$

$$\tag{30}$$

where we have re-labeled M as M^ε to stress the dependence on the small parameter. Note that the exponential-affine term e^{-Bx} does *not* depend on ε. Therefore we will proceed with an asymptotic expansion for M^ε in powers of ε, with the other part appearing as a factor.

To this end, we re-write the equation (24) for M^ε as $\mathcal{A}^\varepsilon M^\varepsilon = 0$, where

$$\mathcal{A}^\varepsilon = \frac{1}{\varepsilon}\mathcal{L}_0 + \frac{1}{\sqrt{\varepsilon}}\mathcal{A}_1 + \mathcal{A}_2, \tag{31}$$

9

and the operators $\mathcal{L}_0, \mathcal{A}_1, \mathcal{A}_2$ are defined as follows:

$$\longrightarrow \qquad \mathcal{L}_0 = \nu^2 \frac{\partial^2}{\partial y^2} + (m - y)\frac{\partial}{\partial y} \,.$$

Note that $\alpha \mathcal{L}_0$ is the infinitesimal generator of the mean-reverting OU process (Y_t) given by (16) under the real-world measure $I\!\!P$.

$$\longrightarrow \qquad \mathcal{A}_1 = -\sqrt{2}\nu b(t, y)\frac{\partial}{\partial y} \,,$$

contains the two market prices of risk and a term due to the correlation ρ.

$$\longrightarrow \qquad \mathcal{A}_2 = \frac{\partial}{\partial t} + c(t, y)1 \,,$$

is the operator associated with the ordinary differential equation which leads to the formula (11) for A in the constant volatility Vasicek problem. Here, it has volatility $f(y)$ and long run mean $(r_\infty - \frac{\lambda(y)f(y)}{a})$ which depend on y.

4.2.3 Asymptotic Expansion and Main Result

The theorem of this section describes the approximation of bond prices we obtain from an asymptotic expansion for M^ε, solution of (24). We introduce the expressions

$$\begin{aligned}
\bar{\sigma}^2 &= \langle f^2 \rangle, \\
r^\star &= r_\infty - \langle \lambda f \rangle / a \qquad\qquad\qquad (32) \\
V_3 &= \frac{\nu\sqrt{\varepsilon}}{\sqrt{2}}\rho\langle f\phi' \rangle \qquad\qquad\qquad (33) \\
V_2 &= -\frac{\nu\sqrt{\varepsilon}}{\sqrt{2}}\langle \Lambda\phi' \rangle - \nu\rho\sqrt{2\varepsilon}\,\langle f\psi' \rangle \\
V_1 &= \nu\sqrt{2\varepsilon}\,\langle \Lambda\psi' \rangle \,,
\end{aligned}$$

where $\langle \cdot \rangle$ denotes the averaging with respect to the density $\Phi(y)$ of the Gaussian distribution $\mathcal{N}(m, \nu^2)$, and ϕ and ψ are specific solutions of the Poisson equations

$$\mathcal{L}_0\phi = f^2 - \langle f^2 \rangle \quad \text{and} \quad \mathcal{L}_0\psi = \lambda f - \langle \lambda f \rangle. \qquad\qquad (34)$$

Indeed, writing $\mathcal{L}_0 = \frac{\nu^2}{\Phi}\frac{\partial}{\partial y}(\Phi\frac{\partial}{\partial y})$, we see that, under the boundedness assumptions on f and λ, we can (and do) choose ϕ and ψ such that their first derivatives given by

$$\phi'(y) = \frac{1}{\nu^2\Phi(y)}\int_{-\infty}^y \left(f(z)^2 - \langle f^2 \rangle\right)\Phi(z)dz\,; \qquad \psi'(y) = \frac{1}{\nu^2\Phi(y)}\int_{-\infty}^y \left(\lambda(z)f(z) - \langle \lambda f \rangle\right)\Phi(z)dz,$$

are bounded and ϕ and ψ themselves are at most linearly growing in $|y|$.

The V's are complicated functions of the model parameters, including market prices of risk, and the volatility function f. Notice that no skew ($\rho = 0$) implies $V_3 = 0$ and a market price of risk such that λf is constant implies $V_1 = 0$.

Further, let

$$\begin{aligned}
D(\tau) &= \frac{V_3}{a^3}\left(\tau - B(\tau) - \frac{1}{2}aB(\tau)^2 - \frac{1}{3}a^2B(\tau)^3\right) \\
&\quad - \frac{V_2}{a^2}\left(\tau - B(\tau) - \frac{1}{2}aB(\tau)^2\right) + \frac{V_1}{a}\left(\tau - B(\tau)\right) \qquad (35) \\
P_0(t, x; T) &= \bar{P}(t, x; T) = A(T - t)e^{-B(T-t)x} \qquad\qquad (36) \\
\tilde{P}_1(t, x; T) &= D(T - t)A(T - t)e^{-B(T-t)x} \qquad\qquad (37)
\end{aligned}$$

10

where A and B were defined in (11) and (10). Notice that P_0 is exactly the Vasicek one-factor bond pricing function $\bar{P}$ with the "averaged" parameters $(a, r^\star, \bar{\sigma})$ related to the stochastic volatility model parameters in (17-18) by (32). The **corrected bond price** is given by

$$P(t, x, y; T) \approx P_0(t, x; T) + \tilde{P}_1(t, x; T) = A(T - t)\left[1 + D(T - t)\right] e^{-B(T-t)x}, \qquad (38)$$

where D is a small factor of order $1/\sqrt{\alpha}$. The error in the approximation (38) is of order $1/\alpha$. This is the main result of this paper stated in the next theorem.

Theorem 1 *Suppose P_0 is given by (36) and $\tilde{P}_1$ by (37). Then for any fixed $t < T$, $x, y \in \mathcal{R}$,*

$$\left| P(t, x, y; T) - \left(P_0(t, x; T) + \tilde{P}_1(t, x; T) \right) \right| = \mathcal{O}(\varepsilon),$$

where $P(t, x, y; T)$ is the model's bond price given by (20).

Proof. From (26), it is enough to show that

$$\left| A(T - t)[1 + D(T - t)] - M^\varepsilon(T - t, y) \right| = \mathcal{O}(\varepsilon).$$

Notice that $A(T - t)$, as a function of t, is the solution of the equation

$$\langle \mathcal{A}_2 \rangle A = \frac{d}{dt} A(T - t) + \bar{c}(t) A(T - t) = 0, \qquad (39)$$

with the terminal condition $A(0) = 1$ at $t = T$, and where

$$\bar{c}(t) = \langle c(t, \cdot) \rangle = \frac{1}{2} \bar{\sigma}^2 B(T - t)^2 - ar^\star B(T - t), \qquad (40)$$

is the average of $c(t, y)$ given by (23). The function $\mathcal{A}_2 A$ being centered (by (39)) and bounded, the Poisson equation in M_2

$$\mathcal{L}_0 M_2 + \mathcal{A}_2 A(T - t) = 0, \qquad (41)$$

admits a solution with bounded first y-derivative. It is given, in terms of the functions ϕ and ψ defined in (34), by

$$M_2(t, y) = -A(T - t)\left(\frac{1}{2}\phi(y) B(T - t)^2 + \psi(y) B(T - t) \right),$$

where we have arbitrarily set the integration constant to zero.

The correction $D(T - t)$ defined by (35) has been chosen such that it is the solution of the equation

$$\langle \mathcal{A}_2 \rangle AD + \sqrt{\varepsilon} \langle \mathcal{A}_1 M_2 \rangle = 0,$$

with the terminal condition $D(0) = 0$ at $t = T$.

Similarly, define $M_3(t, y)$ as a solution of the Poisson equation

$$\mathcal{L}_0 M_3 + \mathcal{A}_2 A(T - t) \frac{D(T - t)}{\sqrt{\varepsilon}} + \mathcal{A}_1 M_2 = 0, \qquad (42)$$

which can also be written in terms of solutions of Poisson equations similar to (34):

$$M_3(t, y) = -A(T - t)\left[\phi_1(y) \frac{D(T - t)}{\sqrt{\varepsilon}} + \nu \sqrt{2}\left(\frac{1}{2}\phi_2(y) B(T - t)^2 + \phi_3(y) B(T - t) \right) \right],$$

11

where ϕ_1, ϕ_2 and ϕ_3 are respectively solutions to:

$$\begin{aligned}
\mathcal{L}_0\phi_1 &= c - \bar{c}, \\
\mathcal{L}_0\phi_2 &= b\phi' - \langle b\phi' \rangle, \\
\mathcal{L}_0\phi_3 &= b\psi' - \langle b\psi' \rangle,
\end{aligned}$$

and $b, c, (\phi, \psi), \bar{c}$ are respectively defined in (22,23,34,40).

Because of the boundedness assumptions on $f(y)$, $\lambda(y)$ and $\gamma(y)$, it follows that the solutions of these Poisson equations can be chosen with bounded first derivatives, and are at most linearly growing in $|y|$. Consequently, M_2 and M_3 are at most linearly growing in $|y|$, and the linear growth rate is uniformly bounded in t since A, B and D are.

Let the error function Z^ε be defined by

$$Z^\varepsilon = A(1+D) + \varepsilon M_2 + \varepsilon^{3/2} M_3 - M^\varepsilon. \tag{43}$$

At the terminal time T we have

$$Z^\varepsilon(T, y) = 0, \tag{44}$$

because of the terminal conditions $M^\varepsilon(T, y) = A(0) = 1$, $D(0) = 0$ and $M_2(T, y) = M_3(T, y) = 0$.

The next step is to compute $\mathcal{A}^\varepsilon Z^\varepsilon$ using the definition (31) of $\mathcal{A}^\varepsilon$, and the properties of $(M^\varepsilon, A, D, M_2, M_3)$, keeping in mind that D is proportional to $\sqrt{\varepsilon}$:

$$\begin{aligned}
\mathcal{A}^\varepsilon Z^\varepsilon &= \mathcal{A}^\varepsilon \left(A(1+D) + \varepsilon M_2 + \varepsilon\sqrt{\varepsilon} M_3 - M^\varepsilon \right) \\
&= \left(\frac{1}{\varepsilon}\mathcal{L}_0 + \frac{1}{\sqrt{\varepsilon}}\mathcal{A}_1 + \mathcal{A}_2 \right) \left(A + \sqrt{\varepsilon}A\frac{D}{\sqrt{\varepsilon}} + \varepsilon M_2 + \varepsilon\sqrt{\varepsilon} M_3 \right) - \mathcal{A}^\varepsilon M^\varepsilon \\
&= \frac{1}{\varepsilon}\mathcal{L}_0 A + \frac{1}{\sqrt{\varepsilon}}(\mathcal{L}_0 A\frac{D}{\sqrt{\varepsilon}} + \mathcal{A}_1 A) \\
&\quad + (\mathcal{L}_0 M_2 + \mathcal{A}_1 A\frac{D}{\sqrt{\varepsilon}} + \mathcal{A}_2 A) + \sqrt{\varepsilon}\left(\mathcal{L}_0 M_3 + \mathcal{A}_1 M_2 + \mathcal{A}_2 A\frac{D}{\sqrt{\varepsilon}} \right) \\
&\quad + \varepsilon \left(\mathcal{A}_1 M_3 + \mathcal{A}_2 M_2 + \sqrt{\varepsilon}\mathcal{A}_2 M_3 \right) \\
&= \varepsilon \left(\mathcal{A}_1 M_3 + \mathcal{A}_2 M_2 \right) + \varepsilon^{3/2}\mathcal{A}_2 M_3, \tag{45}
\end{aligned}$$

because M^ε solves the original equation $\mathcal{A}^\varepsilon M^\varepsilon = 0$ and A, D, M_2 and M_3 have been chosen to cancel the first four terms (the first two terms cancel because A and D do not depend on y, and the third and fourth terms cancel because of (41) and (42)).

Putting together $\mathcal{A}^\varepsilon Z^\varepsilon = \mathcal{O}(\varepsilon)$ and $Z^\varepsilon(T, y) = 0$ from (44), we obtain

$$Z^\varepsilon = \mathcal{O}(\varepsilon). \tag{46}$$

To see this, we denote $\mathcal{A}^\varepsilon Z^\varepsilon = \varepsilon F^\varepsilon$ where

$$F^\varepsilon(t, y) = \mathcal{A}_1 M_3(t, y) + \mathcal{A}_2 M_2(t, y) + \sqrt{\varepsilon}\mathcal{A}_2 M_3(t, y), \tag{47}$$

and we write the probabilistic representation of Z^ε:

$$Z^\varepsilon(t, x, y) = \varepsilon \mathbb{E}^{\star(\lambda,\gamma)}\left\{ -\int_t^T e^{-\int_t^s r_u\,du} F^\varepsilon(s, Y_s)\,ds \mid r_t = x, Y_t = y \right\}. \tag{48}$$

Again, because of our smoothness and boundedness assumptions, F^ε is at most linearly growing in $|y|$. The result (46) follows by using the exponential moments of $\int_t^s r_u\,du$ obtained as a consequence of (28) in the proof of Proposition 1, which are uniformly bounded in ε and t, and the bound on the variance of Y_s given in Appendix B.

Finally, using (43), the conclusion of the proof is given by:

$$|M^\varepsilon - A(1+D)| = |\varepsilon M_2 + \varepsilon^{3/2} M_3 - Z^\varepsilon| \le |Z^\varepsilon| + \varepsilon|M_2 + \sqrt{\varepsilon}M_3| = \mathcal{O}(\varepsilon).$$

12

4.2.4 Consequences

In Figure 2, we show the effect of the small correction $\tilde{P}_1$ on the bond prices (top) and on the
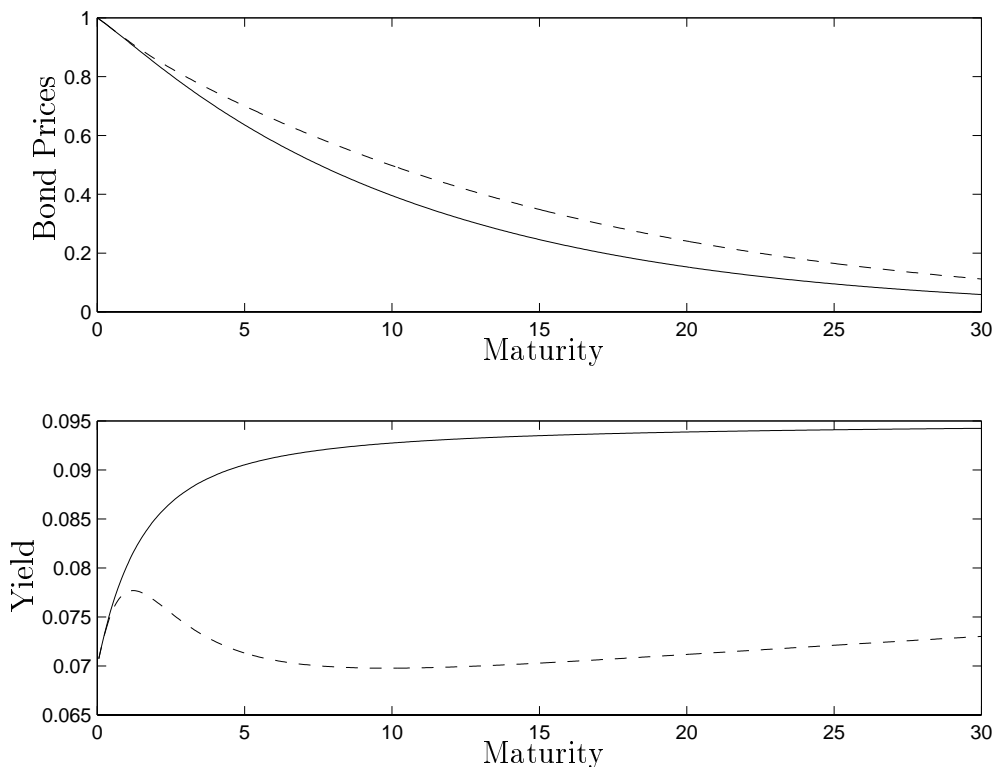

Figure 2: *Top: bond prices and corrected bond prices (dotted curve). Bottom: yield curve and corrected yield curve (dotted curve) in the simulated Vasicek model (constant and stochastic volatility) with: $a = 1$, $r^\star = 0.1$ and $\bar{\sigma} = 0.1$ as in Figure 1. Concerning the correction we have used test parameter values $V_3 = 1/\sqrt{\alpha}$ which assumes a nonzero skew (and in particular would imply from (33) that $\rho\nu\langle f\phi'\rangle = \sqrt{2}$), $\alpha = 10^3$ and $\lambda = \gamma = 0$ implying $V_1 = 0$ and $V_2 = 0$. Maturity runs from 0 to 30 years and the initial rate is $x = 0.07$.*

In the case $\rho = 0$, the effect of the correction is very easy to understand. Using the fact that D' and D are small of order $\sqrt{\varepsilon}$, the following simple calculation

$$
\begin{aligned}
\frac{[A(1+D)]'}{[A(1+D)]} &= \left(\frac{A'}{A} + \frac{D'}{1+D}\right) \\
&= \left(\frac{A'}{A} + D'\right) + \mathcal{O}(\varepsilon) \\
&= \left(\frac{1}{2}\bar{\sigma}^2 B^2 - ar^\star B - V_2 B^2 + V_1 B\right) + \mathcal{O}(\varepsilon) \\
&= \frac{\tilde{A}'}{\tilde{A}} + \mathcal{O}(\varepsilon)
\end{aligned}
$$

shows that

$$
A(1+D) = \tilde{A} + \mathcal{O}(\varepsilon)
$$

13

where $\tilde{A}$ is exactly as in the Vasicek formula (11) with the same terminal condition as $A(1 + D)$, and with the *effective parameters*

$$
\begin{aligned}
\tilde{\sigma}^2 &= \bar{\sigma}^2 - 2V_2 \\
\tilde{r}^\star &= r^\star - \frac{V_1}{a}.
\end{aligned}
$$

In other words, V_2 shifts the level of the long-run average volatility and V_1 contributes an extra risk premium in the drift of the short-rate. These two effects are the result of uncertainty in the volatility.

When $\rho \neq 0$, the part associated to V_3 incorporates more complex effects due to the skew.

4.3   Corrections for Bond Option Prices

We consider now the bond option pricing problem in the context of the stochastic volatility Vasicek model introduced in Section 4.1. The short rate process (r_t) and the OU process (Y_t) driving the volatility are given by equations (17) and (18). The option on the zero-coupon bond is as in Section 3.3, with payoff function $h(L(T_0, T))$. Its price at time t is denoted by $Q(t, x, y; T, T_0)$. It is now a function of the current rate $r_t = x$ and the current volatility level $Y_t = y$, given by

$$
Q(t, x, y; T, T_0) = \mathbb{E}^{\star(\lambda, \gamma)} \left\{ e^{-\int_t^{T_0} r_s \, ds} h\left(L(T_0, T)\right) \mid r_t = x, Y_t = y \right\}, \tag{49}
$$

where the bond price $L(T_0, T)$ is a function of r_{T_0} and Y_{T_0} given by formulas (20) and (27) with t replaced by T_0:

$$
L(T_0, T) = \mathbb{E}^{\star(\lambda, \gamma)} \left\{ e^{-\int_{T_0}^T r_s \, ds} \mid r_{T_0}, Y_{T_0} \right\} = P(T_0, r_{T_0}, Y_{T_0}; T) = M(T_0, Y_{T_0}) e^{-B(T - T_0) r_{T_0}}.
$$

We assume that the payoff function h is at most linearly growing at infinity, and we consider first the case where it is a **smooth function** with bounded derivatives. The nonsmooth case, with the important application to call options on bonds, will be discussed in Section 4.3.3. The exponential moments of r_{T_0} and $\int_t^{T_0} r_s \, ds$ obtained in Proposition 1 enable us to use the Feynman-Kac formula: the premium function $Q(t, x, y; T_0, T)$ is again solution of the partial differential equation (21) but the terminal condition at time T_0 becomes

$$
Q(T_0, x, y; T, T_0) = h\left(P(T_0, x, y; T)\right) = h\left(M(T_0, y) e^{-B(T - T_0) x}\right). \tag{50}
$$

Notice that in (21), the potential "x" is unbounded below. In order to handle this equation, following the forward measure idea of Jamshidian, one can make the substitution

$$
Q(t, x, y; T, T_0) = P(t, x, y; T_0) u(t, y, z),
$$

where

$$
z = \log\left(\frac{P(t, x, y; T)}{P(t, x, y; T_0)}\right),
$$

and easily deduce that $u(t, y, z)$ solves an autonomous linear PDE with time-dependent coefficients and no potential term. This trick gives the regularity properties of u needed to prove that Q solves (21). It turns out that this is not practical for computation and does not simplify the asymptotic analysis presented in this section. This is because the coefficients of the PDE satisfied by $u(t, y, z)$ involve $\frac{\partial M}{\partial y}/M$, where M is given in (24).

14

Observe that the terminal condition depends also on the volatility variable y. Fortunately, in the fast mean-reverting case (α large) we have already obtained, in Section 4.2.3, an expansion of the bond price $P(T_0, x, y; T)$ and we have seen that, up to order $\mathcal{O}(1/\sqrt{\alpha})$, this expansion is independent of y. Using the expansion (38) and Theorem 1 we have:

$$
\begin{aligned}
P(T_0, x, y; T) &= P_0(T_0, x; T) + \tilde{P}_1(T_0, x; T) + \mathcal{O}(\varepsilon) \\
&= A_0(1 + D_0)e^{-B_0 x} + \mathcal{O}(\varepsilon),
\end{aligned}
\tag{51}
$$

where A_0, B_0, D_0 denote respectively $A(T - T_0), B(T - T_0), D(T - T_0)$.

Using the assumed smoothness of the function h, the terminal condition (50) can then be expanded as:

$$
\begin{aligned}
Q(T_0, x, y; T, T_0) &= h(P_0(T_0, x; T)) + \tilde{P}_1(T_0, x; T) h'(P_0(T_0, x; T)) + \mathcal{O}(\varepsilon) \\
&= h\left(A_0 e^{-B_0 x}\right) + A_0 D_0 e^{-B_0 x} h'\left(A_0 e^{-B_0 x}\right) + \mathcal{O}(\varepsilon).
\end{aligned}
\tag{52}
$$

The PDE (21) satisfied by the bond option price function $Q(t, x, y; T, T_0)$, denoted here by $Q^\varepsilon(t, x, y)$, can be rewritten

$$
\mathcal{L}^\varepsilon Q^\varepsilon = \left(\frac{1}{\varepsilon}\mathcal{L}_0 + \frac{1}{\sqrt{\varepsilon}}\mathcal{L}_1 + \mathcal{L}_2\right) Q^\varepsilon = 0,
\tag{53}
$$

where $\mathcal{L}_0$ is the OU infinitesimal generator as in (31), and

$$
\mathcal{L}_2 = \frac{\partial}{\partial t} + \frac{1}{2}f(y)^2 \frac{\partial^2}{\partial x^2} + (a(r_\infty - x) - \lambda(y)f(y))\frac{\partial}{\partial x} - x 1,
$$

is the Vasicek operator with volatility $f(y)$ and long run mean $(r_\infty - \frac{\lambda(y)f(y)}{a})$, and

$$
\mathcal{L}_1 = \sqrt{2}\nu\rho f(y)\frac{\partial^2}{\partial x \partial y} - \sqrt{2}\nu\Lambda(y)\frac{\partial}{\partial y}.
$$

4.3.1   Formal Expansion

Expanding Q^ε as

$$
Q^\varepsilon(t, x, y) = Q_0(t, x, y) + \sqrt{\varepsilon}Q_1(t, x, y) + \varepsilon Q_2(t, x, y) + \cdots,
\tag{54}
$$

where maturities are not shown to simplify the notation, we now carry out the formal derivation of Q_0 and Q_1 while the proof of the accuracy of the approximation will be given in Section 4.3.2.

Zero-order term

Expanding the equation $\mathcal{L}^\varepsilon Q^\varepsilon = 0$, the term of order $1/\varepsilon$, $\mathcal{L}_0 Q_0$, cancels if $Q_0(t, x, y)$ does not depend on y:

$$
Q_0(t, x, y) = Q_0(t, x).
$$

The terms of order $1/\sqrt{\varepsilon}$, $\mathcal{L}_0 Q_1 + \mathcal{L}_1 Q_0 = \mathcal{L}_0 Q_1$, again cancels if $Q_1(t, x, y)$ does not depend on y:

$$
Q_1(t, x, y) = Q_1(t, x).
$$

The $\mathcal{O}(1)$ terms lead to the Poisson equation $\mathcal{L}_2 Q_0 + \mathcal{L}_0 Q_2 = 0$ for Q_2 whose solvability condition, ensuring the existence of solutions in a reasonable space, is $\langle \mathcal{L}_2 Q_0 \rangle = 0$ where $\langle \cdot \rangle$ denotes again

15

the averaging with respect to the invariant distribution of the OU process (Y_t). Since $Q_0(t,x)$ is independent of y, we have

$$\langle \mathcal{L}_2 Q_0 \rangle = \langle \mathcal{L}_2 \rangle Q_0 = \frac{\partial Q_0}{\partial t} + \frac{1}{2}\bar{\sigma}^2 \frac{\partial^2}{\partial x^2} + (a(r_\infty - x) - \langle \lambda f \rangle)\frac{\partial Q_0}{\partial x} - x Q_0 = 0, \qquad (55)$$

with the terminal condition $Q_0(T_0, x) = h\left(P_0(T_0, x; T)\right) = h\left(A_0 e^{-B_0 x}\right)$ at time T_0 which corresponds to the leading order term in the terminal condition (52). This is exactly the problem solved in the previous Section 3.3 in the case of a constant volatility $\bar{\sigma}$ and $r^\star = r_\infty - \langle \lambda f \rangle / a$, as defined in (32). It is the Vasicek one-factor option price which can be written

$$Q_0(t, x) = \mathbb{E}^\star \left\{ e^{-\int_t^{T_0} \bar{r}_s ds} h\left(A_0 e^{-B_0 \bar{r}_{T_0}}\right) \mid \bar{r}_t = x \right\}, \qquad (56)$$

where $(\bar{r}_t)$ is the one-factor process defined by the SDE (4) under the risk-neutral measure $\mathbb{P}^\star$.

First-order correction

The $\mathcal{O}(\sqrt{\varepsilon})$ terms give a Poisson equation in Q_3 whose solvability condition reduces to

$$\langle \mathcal{L}_2 \rangle Q_1 = -\langle \mathcal{L}_1 Q_2 \rangle.$$

Using the same argument as in the proof of Theorem 1, one can write $Q_2 = -\mathcal{L}_0^{-1}(\mathcal{L}_2 - \langle \mathcal{L}_2 \rangle)Q_0$. We set $\tilde{Q}_1(t, x) = \sqrt{\varepsilon} Q_1(t, x)$ and deduce

$$\begin{aligned}
\langle \mathcal{L}_2 \rangle \tilde{Q}_1 &= \sqrt{\varepsilon} \left\langle \mathcal{L}_1 \mathcal{L}_0^{-1}(\mathcal{L}_2 - \langle \mathcal{L}_2 \rangle) \right\rangle Q_0 \\
&= \sqrt{\varepsilon} \left\langle \mathcal{L}_1 \mathcal{L}_0^{-1} \left(\frac{1}{2}(f(y)^2 - \langle f^2 \rangle)\frac{\partial^2}{\partial x^2} - (\lambda(y)f(y) - \langle \lambda f \rangle)\frac{\partial}{\partial x} \right) \right\rangle Q_0 \\
&= \sqrt{\varepsilon} \left\langle \mathcal{L}_1 \left(\frac{1}{2}\phi(y)\frac{\partial^2}{\partial x^2} - \psi(y)\frac{\partial}{\partial x} \right) \right\rangle Q_0 \\
&= \left(V_1 \frac{\partial}{\partial x} + V_2 \frac{\partial^2}{\partial x^2} + V_3 \frac{\partial^3}{\partial x^3} \right) Q_0 \equiv \mathcal{A}Q_0, \qquad (57)
\end{aligned}$$

where the small constants V_1, V_2 and V_3 proportional to $\sqrt{\varepsilon}$ are defined in (33). In other words, $\tilde{Q}_1(t, x)$ satisfies the Vasicek PDE with a small source term given by $\mathcal{A}Q_0(t, x)$, and, using (52), with the terminal condition at time T_0 given by

$$\tilde{Q}_1(T_0, x) = \tilde{P}_1(T_0, x; T)h'\left(P_0(T_0, x; T)\right) = A_0 D_0 e^{-B_0 x} h'\left(A_0 e^{-B_0 x}\right).$$

Using the probabilistic representation of the problems (55) and (57) with their respective terminal conditions, one can write

$$\begin{aligned}
Q(t, x, y; T, T_0) &\approx Q_0(t, x) + \tilde{Q}_1(t, x) \\
&= \mathbb{E}^\star \left\{ e^{-\int_t^{T_0} \bar{r}_s ds} \left[h\left(A_0 e^{-B_0 \bar{r}_{T_0}}\right) + A_0 D_0 e^{-B_0 \bar{r}_{T_0}} h'\left(A_0 e^{-B_0 \bar{r}_{T_0}}\right) \right] \mid \bar{r}_t = x \right\} \\
&\quad - \mathbb{E}^\star \left\{ \int_t^{T_0} e^{-\int_t^u \bar{r}_s ds} \mathcal{A}Q_0(u, \bar{r}_u) du \mid \bar{r}_t = x \right\}. \qquad (58)
\end{aligned}$$

4.3.2 Accuracy of the Bond Option Price Approximation

In this section we prove the following result:

Theorem 2 *Assume the payoff function h is smooth and has bounded derivatives, and suppose Q_0 and $\tilde{Q}_1$ are the classical solutions of the problems (55) and (57) with their respective terminal conditions. Then for any fixed $t < T_0$, $x, y \in \mathcal{R}$,*

$$\left| Q(t, x, y; T, T_0) - \left(Q_0(t, x) + \tilde{Q}_1(t, x) \right) \right| = \mathcal{O}(\varepsilon),$$

where $Q(t, x, y; T, T_0)$ is the model's bond option price given by (49).

Proof. The proof follows the lines of the proof of Theorem 1. It consists again in using the functions $Q_2(t, x, y)$ and $Q_3(t, x, y)$, solutions of the Poisson equations

$$\mathcal{L}_0 Q_2 + \mathcal{L}_2 Q_0 = 0,$$

$$\mathcal{L}_0 Q_3 + \mathcal{L}_1 Q_2 + \mathcal{L}_2 Q_1 = 0,$$

where $Q_0(t, x)$ is the Vasicek option price (56), satisfying $\langle \mathcal{L}_2 Q_0 \rangle = \langle \mathcal{L}_2 \rangle Q_0 = 0$, and $Q_1(t, x)$ has been chosen such that the centering condition $\langle \mathcal{L}_1 Q_2 + \mathcal{L}_2 Q_1 \rangle = 0$ is satisfied. The main difference with the proof of Theorem 1 is that the variables x and y do not separate and therefore we have to control the successive x-derivatives of Q_0 and Q_1 involved in Q_2 and Q_3. This will be done in Lemma 1, at the end of this section, where we will show that these derivatives are at most exponential in x which, combined with the exponential moments of r_t and its time-integrals, will give the conclusion.

As in (43), we introduce the error function

$$\tilde{Z}^\varepsilon = Q_0 + \sqrt{\varepsilon} Q_1 + \varepsilon Q_2 + \varepsilon^{3/2} Q_3 - Q^\varepsilon, \tag{59}$$

with the terminal condition

$$\tilde{Z}^\varepsilon(T, y) = \varepsilon \left(Q_2(T, x, y) + \sqrt{\varepsilon} Q_3(T, x, y) \right) + \mathcal{O}(\varepsilon) \equiv \varepsilon \tilde{G}^\varepsilon(x, y), \tag{60}$$

where we define $\tilde{G}^\varepsilon(x, y) = Q_2(T, x, y) + \sqrt{\varepsilon} Q_3(T, x, y) + \mathcal{O}(1)$, where the $\mathcal{O}(1)$ term is uniformly bounded by the assumptions of smoothness of the payoff h, and boundedness of its derivatives.

Our choice of (Q_0, Q_1, Q_2, Q_3) and a computation similar to (45) gives

$$\mathcal{L}^\varepsilon \tilde{Z}^\varepsilon = \varepsilon \left(\mathcal{L}_1 Q_3 + \mathcal{L}_2 Q_2 \right) + \varepsilon^{3/2} \mathcal{L}_2 Q_3 \equiv \varepsilon \tilde{F}^\varepsilon, \tag{61}$$

where we define $\tilde{F}^\varepsilon(t, x, y) = \mathcal{L}_1 Q_3(t, x, y) + \mathcal{L}_2 Q_2(t, x, y) + \sqrt{\varepsilon} \mathcal{L}_2 Q_3(t, x, y)$.

The error $\tilde{Z}^\varepsilon$ is solution of the problem (61) with the terminal condition (60). The functions Q_2 and Q_3 are linear combinations of the derivatives $\frac{\partial^m Q_0}{\partial x^m}$, for $m \leq 3$, and $\frac{\partial^n Q_1}{\partial x^n}$, for $n \leq 2$, with coefficients which are functions of y, at most linearly growing at infinity. For instance we have

$$Q_2(t, x, y) = \psi(y) \frac{\partial Q_0}{\partial x} - \frac{1}{2} \phi(y) \frac{\partial^2 Q_0}{\partial x^2},$$

with ϕ and ψ introduced in (34). The Poisson equation defining Q_3 leads to a similar formula involving up to the third derivative of Q_0, up to the second derivative of Q_1, and additional functions of y with the same properties as ϕ and ψ. The form of Q_2 and Q_3 is inherited by the functions $\tilde{F}^\varepsilon$ and $\tilde{G}^\varepsilon$ for $m \leq 5$ and $n \leq 4$. We omit this lengthy computation here.

Consequently, by Lemma 1 below, there exist positive constants c_1 and c_2 which depend on T such that
$$|\tilde{F}^\varepsilon|, \, |\tilde{G}^\varepsilon| \leq c_1(1+|y|)e^{-c_2 x},$$
uniform in t, for $0 \leq t < T_0$. The exponential moments of r_t and its time-integrals, obtained in the proof of Proposition 1, enable us to write the probabilistic representation of $\tilde{Z}^\varepsilon$

$$
\begin{aligned}
\tilde{Z}^\varepsilon(t,x,y) \;=\; & \varepsilon I\!\!E^{\star(\lambda,\gamma)}\left\{ e^{-\int_t^T r_s \, ds} \tilde{G}^\varepsilon(r_T, Y_T) \right. \\
& \left. - \int_t^T e^{-\int_t^s r_u \, du} \tilde{F}^\varepsilon(s, r_s, Y_s) \, ds \mid r_t = x, Y_t = y \right\}.
\end{aligned}
\tag{62}
$$

From the uniformity in ε of these exponential moments and of the variance of Y given in Appendix B, we deduce that $|\tilde{Z}^\varepsilon| = \mathcal{O}(\varepsilon)$. The conclusion of the proof follows from (59).

Lemma 1 *Let $Q_0(t,x)$ and $Q_1(t,x)$ be the functions introduced in the previous sections:*

$$
\begin{aligned}
Q_0(t,x) \;=\; & I\!\!E^\star\left\{ e^{-\int_t^{T_0} \bar{r}_s \, ds} h\left(A_0 e^{-B_0 \bar{r}_{T_0}}\right) \mid \bar{r}_t = x \right\}, \\
Q_1(t,x) \;=\; & I\!\!E^\star\left\{ e^{-\int_t^{T_0} \bar{r}_s \, ds} A_0 \tilde{D}_0 e^{-B_0 \bar{r}_{T_0}} h'\left(A_0 e^{-B_0 \bar{r}_{T_0}}\right) \mid \bar{r}_t = x \right\} \\
& - I\!\!E^\star\left\{ \int_t^{T_0} e^{-\int_t^u \bar{r}_s \, ds} \tilde{\mathcal{A}} Q_0(u, \bar{r}_u) du \mid \bar{r}_t = x \right\},
\end{aligned}
$$

where the quantity $\tilde{D}_0 \equiv D_0/\sqrt{\varepsilon}$ and the operator $\tilde{\mathcal{A}} \equiv \mathcal{A}/\sqrt{\varepsilon}$ do not depend on ε.
Then Q_0 and Q_1 have x-derivatives of all orders, and for $\bar{m}$ and $\bar{n}$ fixed integers, there exist positive constants c_1 and c_2 such that
$$\left|\frac{\partial^m Q_0}{\partial x^m}\right|, \, \left|\frac{\partial^n Q_1}{\partial x^n}\right| \leq c_1 e^{-c_2 x},$$
for all $m \leq \bar{m}$, $n \leq \bar{n}$ and $0 \leq t \leq T_0$.

Proof. From (4) we know that the pair $(\bar{r}_u, \int_t^u \bar{r}_s \, ds)$ is, conditionally on $\{\bar{r}_t = x\}$, normally distributed with mean
$$\left(r^\star + (x - r^\star)e^{-a(u-t)}, \; r^\star(u-t) + (x-r^\star)\frac{1 - e^{-a(u-t)}}{a} \right),$$
and covariance
$$
\bar{\sigma}^2 \begin{pmatrix} \frac{1-e^{-2a(u-t)}}{2a} & \int_t^u e^{-a(u-s)}\frac{1-e^{-a(u-s)}}{a} ds \\ \int_t^u e^{-a(u-s)}\frac{1-e^{-a(u-s)}}{a} ds & \int_t^u \left(\frac{1-e^{-a(u-s)}}{a}\right)^2 ds \end{pmatrix}.
$$
We will denote the corresponding Gaussian density by $g_{(t,u)}(z,l)$ where z stands for $\bar{r}_u$ and l stands for $\int_t^u \bar{r}_s ds$. We also denote by $\bar{g}_{(t,u)}(z,l)$ the same density when the x-terms are removed from the means. One can then write

$$
\begin{aligned}
Q_0(t,x) \;=\; & \int\int e^{-l} h\left(A_0 e^{-B_0 z}\right) g_{(t,T_0)}(z,l) \, dz \, dl \\
\;=\; & \int\int e^{-\left(l + x\frac{1-e^{-a(T_0-t)}}{a}\right)} h\left(A_0 e^{-B_0\left(z + xe^{-a(T_0-t)}\right)}\right) \bar{g}_{(t,T_0)}(z,l) \, dz \, dl.
\end{aligned}
$$

Taking successive derivatives with respect to x by the smoothness of h and the boundedness of its derivatives and dominated convergence due to exponential moments of $\bar{g}_{(t,T_0)}$ gives the part of

18

Lemma 1 concerning Q_0.
Similarly one can write

$$Q_1(t,x) =$$

$$\int\int e^{-\left(l+x\frac{1-e^{-a(T_0-t)}}{a}\right)} A_0 \tilde{D}_0 e^{-B_0\left(z+xe^{-a(T_0-t)}\right)} h'\left(A_0 e^{-B_0\left(z+xe^{-a(T_0-t)}\right)}\right) \bar{g}_{(t,T_0)}(z,l)dz\,dl$$

$$-\int_t^{T_0}\left(\int\int e^{-\left(l+x\frac{1-e^{-a(u-t)}}{a}\right)}(\tilde{\mathcal{A}}Q_0)(u,z+xe^{-a(u-t)})\bar{g}_{(t,u)}(z,l)dz\,dl\right)du\,.$$

The conclusion is as for Q_0, except for the second term, where we use the result for Q_0 and the uniformity with respect to u, between t and T_0, of the exponential moments of $\bar{g}_{(t,u)}$.

4.3.3   Non-smooth payoffs: Call Options on Bonds

The case of non-smooth payoff functions can be treated by a nontrivial regularization argument which is presented in detail in a separate work in the context of equity call options in (Fouque, Papanicolaou, Sircar and Solna 2003).

The argument consists of regularizing the payoff h. Denoting by h^δ this regularized payoff, with $\delta > 0$ a small parameter such that $|h - h^\delta| = \mathcal{O}(\delta)$, we introduce the corresponding quantities $Q^{\varepsilon,\delta}, Q_0^\delta$ and $\tilde{Q}_1^\delta$ analogous to Q^ε, Q_0 and $\tilde{Q}_1$ with h replacing h^δ. We have

$$|Q^\varepsilon - (Q_0 + \tilde{Q}_1)| \leq |Q^\varepsilon - Q^{\varepsilon,\delta}| + |Q^{\varepsilon,\delta} - (Q_0^\delta + \tilde{Q}_1^\delta)| + |Q_0^\delta - Q_0| + |\tilde{Q}_1^\delta - \tilde{Q}_1|.$$

The first term $|Q^\varepsilon - Q^{\varepsilon,\delta}|$ is of order $\mathcal{O}(\delta)$ and we choose $\delta = \varepsilon$. The second term can be shown to be of order $\mathcal{O}(\varepsilon)$ by going through the proof of Theorem 2 with smooth payoff h^δ and $\delta = \varepsilon$. The last two terms are of order $\delta = \varepsilon$ by straightforward computations. In all these estimates, the existence of exponential moments is crucial. Note also that the situation here is much simpler than in the case of equity options presented in (Fouque et al. 2003). This is because taking x-derivatives of $h^\delta(A_0 e^{-B_0 x})$ does not introduce any singularity at $t = T_0$, in contrast to derivatives (with respect to the stock price) of $e^{-d_1^2/2}$ in the Black-Scholes case with the usual notation.

For instance, in the case of a European call bond option, the first term $Q_0(t,x)$ is equal to $\bar{C}(t,x)$ obtained explicitly in the constant volatility case (12-13-14) and the correction is

$$\tilde{Q}_1(t,x) = \mathbb{E}^\star\left\{e^{-\int_t^{T_0}\bar{r}_s ds} A_0 D_0 e^{-B_0\bar{r}_{T_0}} 1_{\{A_0 e^{-B_0\bar{r}_{T_o}}>K\}}|\bar{r}_t = x\right\}$$

$$- \mathbb{E}^\star\left\{\int_t^{T_0} e^{-\int_t^u \bar{r}_s ds}\mathcal{A}\bar{C}(u,\bar{r}_u)du|\bar{r}_t = x\right\}\,.$$

Using the first term in formula (12) for $\bar{C}(t,x)$, the corrected call bond price becomes

$$\bar{C}(t,x) + D_0\bar{P}(t,x;T)N(h_1) - \int_t^{T_0}\mathbb{E}^\star\left\{e^{-\int_t^u \bar{r}_s ds}\mathcal{A}\bar{C}(u,\bar{r}_u)|\bar{r}_t = x\right\}du\,,$$

where the bond price $\bar{P}(t,x;T)$ is given by (7), $D_0 = D(T - T_0)$ is given by (35), and the operator $\mathcal{A}$ is given by (57).

4.4   Implications

The main implication of the asymptotic calculation is reduction of the parametric dependence of the pricing formulas on a specific model. In the table below, the left column lists the original model parameters, including the risk premium functions and the right column the group parameters that need to be estimated from the yield curve using formula (38).

19

Model Parameters	Correction Parameters
Rate of mean-reversion of short-rate a	a
Long-run mean under $I\!P$, r_∞	$r^\star$
Specific volatility function $f(\cdot)$	Effective volatility $\bar\sigma$
Rate of mean-reversion of volatility α	Group parameter V_1
Mean-level of (Y_t) m	Group parameter V_2
"V-vol" β	Group parameter V_3
Correlation ρ	
Interest-rate risk premium $\lambda(\cdot)$	
Volatility risk premium $\gamma(\cdot)$	

In particular, the sensitivities to the correlation ρ and the rate of mean-reversion α of the *hidden* process (Y_t) are absorbed in the group parameters V_1, V_2 and V_3. There is also the added benefit that the function $f(\cdot)$, $\lambda(\cdot)$ and $\gamma(\cdot)$ do not have to be specified, and that to this level of approximation, the present level y of the unobserved volatility-driving process does not need to be estimated.

The stability of the group parameters over time estimated from data is studied in (Cotton 2001).

The stochastic volatility-corrected prices of bond options also depend on these same parameters in the right column to this order. That is, although the group parameters do not give us enough to recover the law of the processes (r_t, Y_t), they do give us enough to price other interest-rate derivatives under fast mean-reverting stochastic volatility. We refer to the previous work on equity derivatives described at the end of Section 2.2 for details about barrier, Asian and American contracts.

5 Illustration from Data

The short rate model with stochastic volatility correction can be tested in a variety of real world (measure) or risk neutral settings. We find examples from each of these two domains (respectively, the modelling of short rate dynamics and contemporaneous yield curve fitting) where a separation of time scales is suggested by our preliminary data analysis.

5.1 Analysis of Interest Rate Dynamics

In this section we examine the dynamics of empirical short term yields. Using a test described in the following section we present evidence for rapid oscillations in short rate volatility on a time scale much shorter than the typical length of bond and bond option contracts. The latter is on the order of years, the former on the order of weeks.

5.1.1 The Method of Variogram Analysis

Variogram analysis can be used in the analysis of any time series. The "ExpOU" stochastic volatility short rate model, obtained by using the exponential function $f(y) = e^y$ in equation (17), is convenient because it gives an explicit function for the variogram for illustration purposes. We model the normalized short rate fluctuations
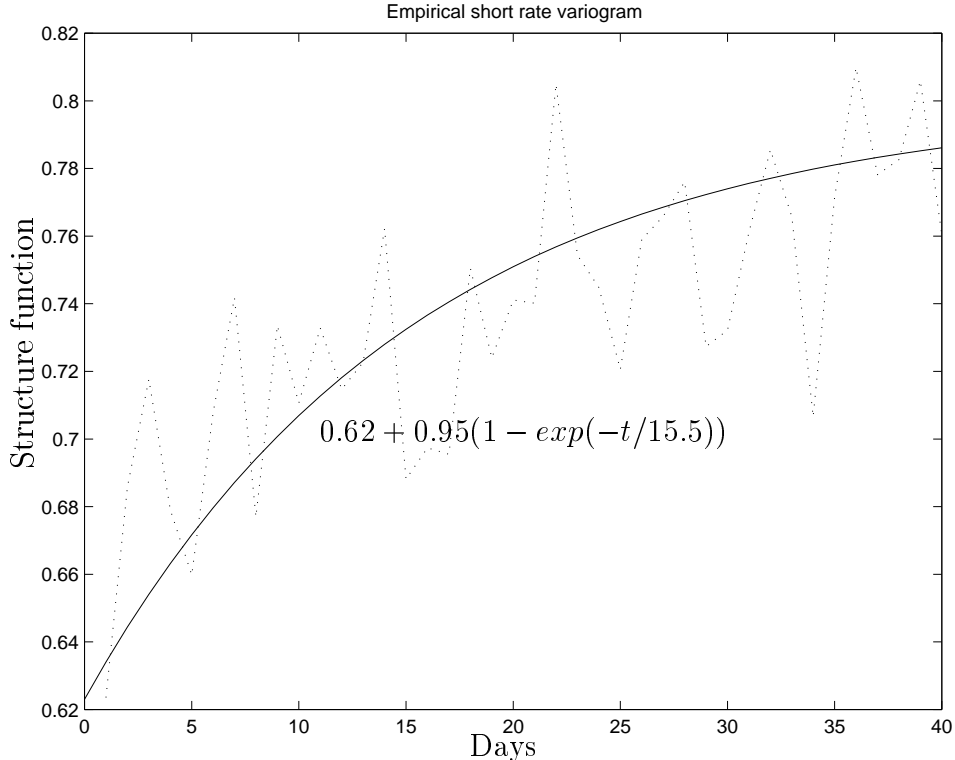
$$D_j = \frac{x_j - x_{j-1}}{\sqrt{\Delta t}}$$

$$0.62 + 0.95(1 - exp(-t/15.5))$$

Figure 3: *Variogram using the logarithm of the absolute value of the normalized differenced US short rate, illustrating evidence for a fast mean reverting component in short rate volatility. The x-axis represents the lag or window length j (in days). The y-axis represents the structure function as defined in the text. The dotted line represents the numerical results for V_j as defined in the text. The solid line is a least squares fitted theoretical prediction based on the ExpOU SV model and given by equation (63). The fitted curve implies a decorrelation length of around fifteen and a half days. The standard deviation of decorrelation time estimates in bootstrapped samples is around four days.*

as

$$D_j = e^{y_j} \epsilon_j$$

where ϵ_j are independent Gaussian random variables with mean zero and variance one. Here x_j is the short rate proxy for the j'th trading day and Δt is one day. We denote the value of the volatility driving process Y_t on the j'th trading day by y_j. This description is consistent with the ExpOU stochastic volatility model up to a short rate drift term which is small when Δt is small. We discard this term for simplicity as it makes negligible difference to the results which follow. We also ignore for purposes of data analysis the cutoff at 0 and ∞ that would have to be applied to the exponential function to satisfy the hypotheses of Girsanov's theorem.

We examine the **empirical structure function** or **variogram** of the log absolute value of the normalized fluctuations. That is, we examine

$$V_j := \frac{1}{N} \sum_n \left(X_{n+j} - X_n \right)^2$$

where

$$X_j := \log |D_j|.$$

21

According to our model,

$$X_j = y_j + \log |\epsilon_j|.$$

To avoid the singularities from the noise $\log |\epsilon_j|$ we pass the X_j through a median filter. The graph of V_j as a function of j is fitted (by least squares) to the theoretical curve

$$\tilde{V}_j = 2\psi^2 + 2\frac{\beta^2}{2\alpha}(1 - \exp(-j/l)), \qquad j \geq 1, \tag{63}$$

which is derived by direct calculation from the OU model for (Y_t). The constants ψ^2, $\frac{\beta^2}{2\alpha}$ and l are the variance of the noise $\log |\epsilon_j|$, the variance of the volatility process y_j, and the decorrelation length of the volatility process y_j respectively. We are most interested in the latter, since it gives α via $l = 1/(\alpha \Delta t)$. We have also assumed here $\rho = 0$. Extensive simulations (in the case of a Black-Scholes model with stochastic volatility) in (Fouque, Papanicolaou, Sircar and Solna 2002) show that the variogram properties are quite insensitive to ρ.

We remark that a similar methodology could be applied to other stochastic volatility models in the class described in this paper (that is, different choices of f), for which the variogram will exhibit similar exponential decay.

5.1.2 Variograms for US Interest Rates

Figure 3 shows the empirical structure function for a time series proxying the US short rate, reflecting the short end of the yield curve and cash deposit rates[1].

The fitted curve implies a decorrelation length of around <u>fifteen and a half days</u>. Short rate simulation was used to test the effect of the median filter and bootstrap errors in the decorrelation time. The standard deviation of decorrelation time estimates in bootstrapped samples was around four days. Figure 4 shows a fitted variogram using one path of simulated data, treated in identical fashion to the real data.

The rough sampling error estimate, arrived at by fitting several hundred structure functions for different simulated data paths, demonstrates the difficulty of measuring a precise time scale of mean reversion in the unobserved volatility driving process. However, this is unimportant. The observation that volatility mean reverts (on the order of one and a half to two and a half weeks) rather than years is sufficient justification for our asymptotic approach.

Further discussion on the variogram technique for validating the existence of a rapidly mean reverting component of short rate volatility is described in detail in (Cotton 2001) and has also been applied to $S\&P500$ index data in (Fouque et al. 2002). In the latter, we demonstrate by simulation the insensitivity of the order estimate to the zero correlation assumption and the median filtering.

5.2 Contemporaneous Yield Curve Fitting

It is desired in many applications to have a model that is flexible enough to fit a wide variety of yield curves, is consistent with a dynamic for the short rate, and yet remains parsimonious. We contend that the Vasicek model with stochastic volatility corrections is a step in this direction.

We examine daily zero coupon yield curves constructed from a variety of bills and USD/Libor swap rates sampled at sixty maturities ranging from one month to thirty years. We compare fitting the yield curve day-by-day to two models: our two-term asymptotic approximation and the CIR model with Poisson down jumps. In both cases, we fixed the level of the short rate r_t to that

---

[1]We are grateful to Dmitri Kachin, formerly of Morgan Stanley, for assistance in this regard.
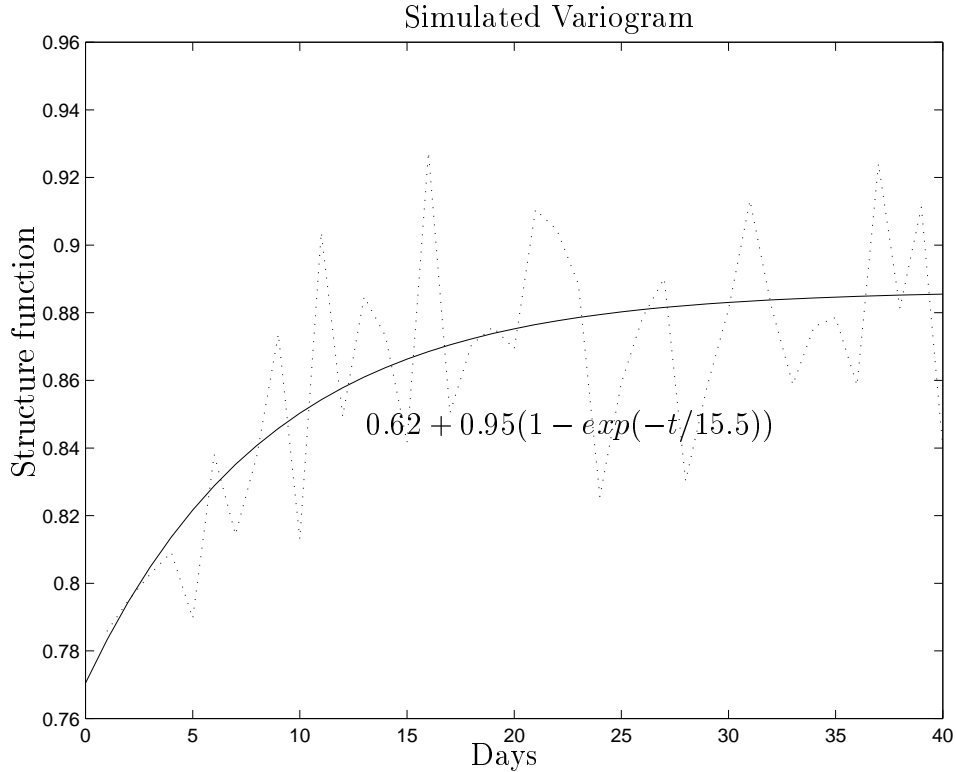
Figure 4: *Variogram for simulations from an ExpOU SV Vasicek model. The parameters used in this simulation were $\bar{r}_\infty = 0.06$, $\alpha = 26$, $m = -5$, $\beta = 2$, $\rho = 0.6$ and $\lambda = \gamma = 0$. Numerical results are dotted. The solid line is a least squares fit for the theoretical prediction as explained in the text. The agreement with the data (compare figure 3) is excellent, insofar as the qualitative shape of the curve is concerned. Note that the* level *of the curve is unimportant for estimating α, and rather reflects the noise driving the short rate process and not the volatility fluctuations - see equation (63). A flat curve, as observed in the variograms of data simulated using multifactor CIR or Vasicek processes, would indicate a small α. Here the qualitative upward curving trend in the data points suggests fast mean reversion.*

proxied from the short end of the curve. This was done for every day in a three year period (ending in 1999).

From the asymptotic theory, there are six parameters $(a, r^\star, \bar{\sigma}, V_1, V_2, V_3)$ to fit, the last three representing effects from changing volatility. To make a fair comparison, we fixed two of these, $\bar{\sigma}$ and a, and fitted the remaining four so that this calibration had less than the five degrees of freedom of the CIR with jumps model. (These two were chosen after experimenting with other choices so that the reduced calibration gave generally good fits of the data and was not overly handicapped by the imposed restriction). Comparison with other models, for example two-factor CIR is a subject for future investigation, so the present results allow us to compare, informally, the alternatives of modifying a simple one-factor model with an asymptotic approximation for stochastic volatility against modifying a one-factor model with jumps.

It was found that excellent fits could be obtained for all days in the sample with only four free parameters. The flexibility of the stochastic volatility model was clearly superior to competing single factor models with four or five free parameters during periods of unusual yield curve shapes.

A snapshot comparison of the yield curve fits obtained with the corrected Vasicek model and

a CIR model with jumps respectively, is shown in Figure 5. Note that the CIR model with jumps, fitted in the same manner as the approximated stochastic volatility model (i.e., fitted contemporaneously, with the short rate determined by the short end of the yield curve), had five free parameters but performed significantly worse, across the thousand or so days on which contemporaneous fits were carried out.
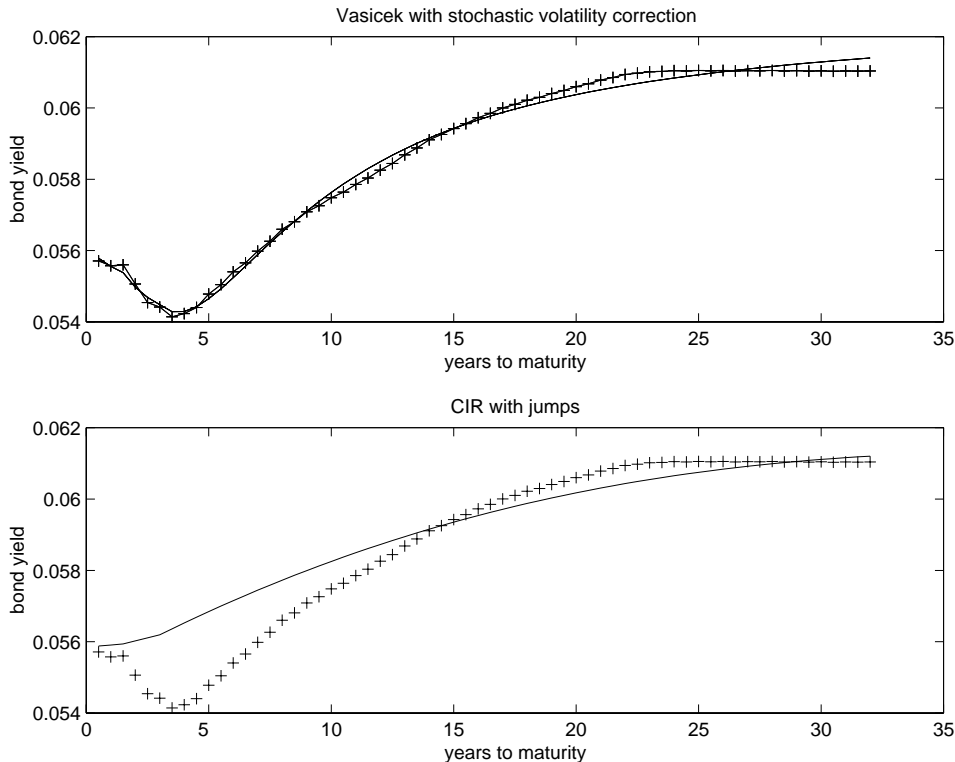

Figure 5: *Snapshot of the yield curve fit with the stochastic volatility corrected Vasicek model (top) and with the single factor CIR model and down jumps (bottom) for September 6, 1998.*

We suggest that the scale separation implicit in the yield formulas (with corrections) greatly assists the fitting of unusual yield curves (as in the crisis of the summer of 1998) without sacrificing parsimony. With the exception of three days in the sample of over one thousand, the root mean square (across maturities) yield curve fitting errors for the corrected Vasicek model were less than six basis points, with the average around two basis points. In contrast, single factor Vasicek and CIR models typically misfit the short to medium yields *during the crisis period* by twenty or more basis points (as in Figure 5). The relative bond pricing error, which is approximately the yield error multiplied by the maturity, is on the order of 0.6% for the 30-year zero coupon bond, using the corrected Vasicek model.

A   Asymptotics around CIR model

We give a heuristic derivation of the fast mean-reverting stochastic volatility asymptotic calculation for a model built around the popular CIR model to illustrate the flexibility of the method under changes to the background short-rate model. Because we would like a closed-form correction for bond prices, we mimic the form of the two-factor CIR model (1-2) in which the independent

Brownian motions have the same square-root-of-affine coefficients in front of them. This means that the derivative pricing PDE has linear coefficients which admits an affine yield curve solution.

We shall look at the model

$$
\begin{aligned}
dr_t &= a(r^\star - r_t)dt + f(Y_t)\sqrt{r_t}\,dW_t^\star, \\
dY_t &= \alpha r_t(m - Y_t)dt + \beta\sqrt{r_t}\left(\rho\,dW_t^\star + \rho'\,dZ_t^\star\right),
\end{aligned}
$$

where we are already under the pricing measure. We will not look at the link between the real measure and the risk-neutral measure for this model. Notice that the second process (Y_t) driving the volatility is no longer autonomous in that it depends on (r_t). We make this change to admit yield curves that are affine in the short-rate level $x = r_t$, but not of course in $y = Y_t$. This allows a closed-form (up to solution of ODEs) for the stochastic volatility correction. We assume $f(y)^2 < 2ar^\star$ for all y, which guarantees existence of a strong solution for (r_t).

The setup for bond pricing is similar to the formal calculation presented in Section 4.2 with the new differential operators

$$
\begin{aligned}
\mathcal{L}_0 &= x\left(\nu^2\frac{\partial^2}{\partial y^2} + (m - y)\frac{\partial}{\partial y}\right) \\
\mathcal{L}_1 &= \sqrt{2}\nu\rho x f(y)\frac{\partial^2}{\partial x\partial y} \\
\mathcal{L}_2 &= \frac{\partial}{\partial t} + \frac{1}{2}f(y)^2 x\frac{\partial^2}{\partial x^2} + a(r^\star - x)\frac{\partial}{\partial x} - x1.
\end{aligned}
$$

Notice that $\mathcal{L}_0$ is merely x times the generator of the OU process (Y_t) scaled by $1/\alpha$, and that $\mathcal{L}_2$ is the CIR operator with volatility $f(y)$.

Then the arguments of Section 4.2 go through analogously, with the extra x-factor in $\mathcal{L}_0$ not affecting the conclusions following from the Poisson equations in y. Bond prices $P^\varepsilon(t, x, y)$ are approximated by

$$
P^\varepsilon(t, x, y) \approx P_0(t, x) + \tilde{P}_1(t, x),
$$

where the first-term solves

$$
\begin{aligned}
\langle\mathcal{L}_2\rangle P_0 = \left(\frac{\partial}{\partial t} + \frac{1}{2}\bar{\sigma}^2 x\frac{\partial^2}{\partial x^2} + a(r^\star - x)\frac{\partial}{\partial x} - x\right)P_0 &= 0 \\
P_0(T, x) &= 1,
\end{aligned}
$$

and is given by the one-factor affine CIR formula

$$
\begin{aligned}
P_0(t, x) &= A(T - t)e^{-B(T-t)x} \\
A(\tau) &= \left(\frac{2\theta e^{(\theta + a)\tau/2}}{(\theta + a)(e^{\theta\tau} - 1) + 2\theta}\right)^{2ar^\star/\bar{\sigma}^2} \tag{64} \\
B(\tau) &= \frac{2(e^{\theta\tau} - 1)}{(\theta + a)(e^{\theta\tau} - 1) + 2\theta} \tag{65} \\
\theta &= \sqrt{a^2 + 2\bar{\sigma}^2}.
\end{aligned}
$$

The correction can be shown to satisfy

$$
\begin{aligned}
\langle\mathcal{L}_2\rangle\tilde{P}_1 &= V_3 x\frac{\partial^3 P_0}{\partial x^3}, \\
P_1(T, x) &= 0,
\end{aligned}
$$

25

where V_3 is some group parameter related to the original parameters that will be estimated from the yield curve. This can be solved to give

$$\tilde{P}_1(t, x) = (D_1(T - t)x + D_2(T - t)) A(T - t)e^{-B(T-t)x},$$

where A and B are given by (64-65) and D_1 and D_2 satisfy the ODEs

$$
\begin{aligned}
D_1' &= V_3 B^3 - (\bar{\sigma}^2 B + a)D_1 \\
D_2' &= ar^\star D_1,
\end{aligned}
$$

with zero initial conditions.

B   Estimate on Moments of Y

In this appendix we derive the estimate on the second moment of Y_s needed at the end of the proof of Theorem 1, and Theorem 2 as well. In fact we do more and derive exponential moments estimates. Recall that the process Y is the solution of the stochastic differential equation

$$dY_s = \frac{1}{\varepsilon}\left(m - Y_s\right)ds - \frac{\nu\sqrt{2}}{\sqrt{\varepsilon}}\Lambda(Y_s)ds + \frac{\nu\sqrt{2}}{\sqrt{\varepsilon}}d\hat{Z}_s^\star, \tag{66}$$

where $\hat{Z}^\star = \rho W^\star + \rho' Z^\star$ is a standard $I\!\!P^{\star(\lambda,\gamma)}$-Brownian motion, $\Lambda = \rho\lambda + \rho'\gamma$ is bounded, and $t \leq s \leq T$ with $Y_t = y$ as in the probabilistic representation of Z^ε in (48).

We start by exponential moments estimates under the real world probability measure $I\!\!P$ for which

$$dY_s = \frac{1}{\varepsilon}\left(m - Y_s\right)ds + \frac{\nu\sqrt{2}}{\sqrt{\varepsilon}}d\hat{Z}_s, \tag{67}$$

where $\hat{Z} = \rho W + \rho' Z$ is a standard $I\!\!P$-Brownian motion, and $t \leq s \leq T$ with $Y_t = y$. In the following, we use the shorthand $I\!\!E_{t,y}\{\cdot\}$ for $I\!\!E\{\cdot \mid Y_t = y\}$.

Lemma 2 *There exists a constant $c(v, y)$ such that for any $t \leq s \leq T$ and $\varepsilon > 0$ we have*

$$I\!\!E_{t,y}\{e^{vY_s}\} \leq c(v, y). \tag{68}$$

Proof. Under $I\!\!P$, the process Y is an OU process converging exponentially fast to its invariant distribution. Indeed Y_s is Gaussian with mean $m + (y - m)e^{-(s-t)/\varepsilon}$ and variance $\nu^2(1 - e^{-2(s-t)/\varepsilon})$. Consequently

$$I\!\!E_{t,y}\{e^{vY_s}\} = \exp\left\{v\left(m + (y - m)e^{-(s-t)/\varepsilon}\right) + \frac{v^2\nu^2}{2}\left(1 - e^{-2(s-t)/\varepsilon}\right)\right\} \tag{69}$$

and the proof follows by choosing $c(v, y) = \exp\left(|v|(|m| + |y - m|) + v^2\nu^2/2\right)$.

We now come back to Y under the risk-neutral probability $I\!\!P^{\star(\lambda,\gamma)}$ denoted by $I\!\!P^\star$ for simplicity.

Proposition 2 *There exists a constant $C(t, T, v, y)$ such that for any $t \leq s \leq T$ and $\varepsilon > 0$ we have*

$$I\!\!E^\star_{t,y}\{e^{vY_s}\} \leq C(t, T, v, y). \tag{70}$$

Proof. By Girsanov's theorem, we have

$$\mathbb{E}^{\star}{}_{t,y}\{e^{vY_s}\} = \mathbb{E}_{t,y}\{e^{vY_s} M_s^{(\Lambda)}\} \tag{71}$$

where $M^{(\Lambda)}$ defined by

$$M_s^{(\Lambda)} = e^{-\int_t^s \Lambda(Y_u)d\hat{Z}_u - \frac{1}{2}\int_t^s \Lambda(Y_u)^2 du}, \tag{72}$$

is a $(\mathbb{P}, (\mathcal{F}_t))$-martingale, and $\hat{Z}$ is a $\mathbb{P}$-standard Brownian motion. We rewrite (71) as

$$\mathbb{E}_{t,y}\{e^{vY_s} M_s^{(\Lambda)}\} = \mathbb{E}_{t,y}\left\{ e^{vY_s} e^{\frac{1}{2}\int_t^s \Lambda(Y_u)^2 du} \sqrt{M_s^{(2\Lambda)}} \right\}, \tag{73}$$

and, using the Cauchy-Schwarz inequality, we deduce that

$$\mathbb{E}^{\star}{}_{t,y}\{e^{vY_s}\} \leq \sqrt{\mathbb{E}_{t,y}\left\{ e^{2vY_s} e^{\int_t^s \Lambda(Y_u)^2 du} \right\}}, \tag{74}$$

since $M^{(2\Lambda)}$ is a martingale with expected value equal to one. Therefore

$$\mathbb{E}^{\star}{}_{t,y}\{e^{vY_s}\} \leq e^{\frac{1}{2}(s-t)||\Lambda||_\infty^2} \sqrt{\mathbb{E}_{t,y}\left\{ e^{2vY_s} \right\}} \tag{75}$$

$$\leq e^{\frac{1}{2}(s-t)||\Lambda||_\infty^2} \sqrt{c(2v, y)}, \tag{76}$$

where $c(v, y)$ is the constant obtained in Lemma 2. We conclude by setting

$$C(t, T, v, y) = e^{\frac{1}{2}(T-t)||\lambda||_\infty^2} \sqrt{c(2v, y)}. \tag{77}$$

Acknowledgment

References

Cotton, P. (2001). *An analytic approach to Ornstein-Uhlenbeck processes with fluctuating parameters and applications in the modeling of fixed income securities*, PhD thesis, Stanford University.

Duffie, D. (2001). *Dynamic Asset Pricing Theory*, 3rd edn, Princeton University Press.

Duffie, D. and Kan, R. (1996). A yield-factor model of interest rates, *Math. Finance* **6**(4): 379–406.

Fouque, J.-P., Papanicolaou, G. and Sircar, K. (1999). Financial modeling in a fast mean-reverting stochastic volatility environment, *Asia-Pacific Financial Markets* **6**(1): 37–48.

Fouque, J.-P., Papanicolaou, G. and Sircar, K. (2000a). Calibrating random volatility, *RISK* **13**(2): 89–92.

Fouque, J.-P., Papanicolaou, G. and Sircar, K. (2001). From the implied volatility skew to a robust correction to Black-Scholes American option prices, *International Journal of Theoretical and Applied Finance* **4**(4): 651–75.

Fouque, J.-P., Papanicolaou, G. and Sircar, R. (2000b). Mean-Reverting Stochastic Volatility, *International Journal of Theoretical and Applied Finance* **3**(1): 101–42.

Fouque, J.-P., Papanicolaou, G., Sircar, K. and Solna, K. (2002). Short time-scale in S&P 500 volatility, *J. Computational Finance.* To appear.

Fouque, J.-P., Papanicolaou, G., Sircar, K. and Solna, K. (2003). Singular perturbations in option pricing, *SIAM J. Applied Math.* To appear.

Frey, R. (1996). Derivative asset analysis in models with level-dependent and stochastic volatility, *CWI Quarterly* **10**(1): 1–34.

Hull, J. (1999). *Options, Futures and Other Derivative Securities*, 4th edn, Prentice Hall.

Jamshidian, F. (1989). An exact bond option pricing formula, *J. Finance* **44**: 205–9.

Krylov, N. V. (1980). *Controlled Diffusion Processes*, Springer-Verlag.

Longstaff, F. and Schwartz, E. (1992). Interest rate volatility and the term structure: a two-factor general equilibrium model, *J. Finance* **XLVII**(4): 1259–1282.

Rogers, L. C. G. (1995). Which model for the term-structure of interest rates should one use?, *Mathematical Finance*, Vol. 65 of *IMA*, Springer NY, pp. 93–116.

Vasicek, O. (1977). An equilibrium charcaterization of the term structure, *J. Financial Econ.* **5**: 177–188.