A Family Of Term-structure Models with Stochastic Volatility

Andrew J.G. Cairns
and
Samuel A. Garcia Rosas

Actuarial Mathematics and Statistics,
School of Mathematical and Computer Sciences,
Heriot-Watt University,
Edinburgh, EH14 4AS, United Kingdom
E-mail: A.Cairns@ma.hw.ac.uk
Internet: www.ma.hw.ac.uk/∼andrewc/
First version: December 15, 2003
This version: August 11, 2004

Abstract

In this paper we extend the class of multifactor term-structure models proposed by Cairns (2004a) to incorporate a more explicit form of stochastic volatility. The models are built up within the framework proposed by Flesaker & Hughston (1996).

Our general aim is to work with models in which zero-coupon bond prices can be expressed in the form

$$P(t,T) = \frac{\int_{T-t}^{\infty} e^{A(u)+B(u)^T X(t)} du}{\int_{0}^{\infty} e^{A(u)+B(u)^T X(t)} du}$$

for some n-dimensional, stationary diffusion $X(t)$ and for suitable deterministic functions $A(u)$ and $B(u)$. We prove that the models require a multivariate affine state-variable $X(t)$ as developed previously by Duffie & Kan (1996). The remainder of the paper describes some numerical experiments for specific two and three-factor models which incorporate one stochastic volatility component.

The models have a close relationship with recently developed market models incorporating stochastic volatility. The new models can therefore be used to provide practitioners with a parsimonious benchmark against which more elaborate market models can be compared.

Keywords: term-structure model; multifactor; positive interest; stochastic volatility; time-homogeneous; log-normal; term-structure of volatility.

1 Introduction

In this paper we will build on a recent work by Cairns (2004a) who proposed a new family of time-homogeneous, multifactor, term-structure models using the Flesaker & Hughston (1996) framework.[1] In this section we will review briefly the earlier paper, before introducing some possible extensions alluded to in the final section of Cairns (2004a). In Section 2 we will provide the theoretical basis for what we will call the Integrated Affine class of term-structure models. This draws heavily on the work of Duffie & Kan (1996). In Sections 3 and 4 we will investigate the characteristics of two models which can be regarded as extensions of Cairns (2004a) models to include stochastic volatility.

Let $P(t,T)$ be the price at t for a zero-coupon bond which matures at time T. Cairns (2004a) developed Flesaker-Hughston models for prices which could be expressed in the form

$$P(t,T) \;=\; \frac{\int_{T-t}^{\infty} H(u,X(t))du}{\int_0^{\infty} H(u,X(t))du} \tag{1.1}$$

$$\text{where}\;\; H(u,x) \;=\; \exp\left[-\beta u + \sum_{i=1}^{n} \sigma_i x_i e^{-\alpha_i u} - \frac{1}{2}\sum_{i,j=1}^{n} \frac{\rho_{ij}\sigma_i\sigma_j}{\alpha_i + \alpha_j} e^{-(\alpha_i+\alpha_j)u}\right] \tag{1.2}$$

In these equations the state variable $X(t) = (X_1(t), \ldots, X_n(t))'$ is an n-dimensional vector of correlated Ornstein-Uhlenbeck processes. The dynamics of the $X_i(t)$ are dependent on a standard n-dimentional Brownian motion $\hat{Z}(t)$ under a pricing measure $\hat{P}$, and are governed by the stochastic differential equations, for $i = 1, \ldots, n$,

$$dX_i(t) \;=\; -\alpha_i X_i(t)dt + \sum_{j=1}^{n} c_{ij} d\hat{Z}_j(t).$$

The pricing measure $\hat{P}$ could be different from the real-world measure P and the traditional risk-neutral measure Q. Without loss of generality (since we can rescale the σ_i parameters) we can impose the requirement that $\sum_{j=1}^{n} c_{ij}^2 = 1$ for all i. If we then define the matrix $C = (c_{ij})_{i,j=1}^n$, CC' represents the instantaneous correlation matrix for the n processes $X_1(t), \ldots, X_n(t)$. This class of model was found to satisfy certain desirable features required of models for use in long-term risk management:

1. all interest rates remain strictly positive;

2. all interest rates can get arbitrarily close to zero;

3. the model is mean reverting;

---

[1] The framework was subsequently generalised by Rogers (1997) and Rutkowski (1997). A summary of these frameworks can be found in Cairns (2004b).

4. suitably parametrised, the model should give rise to long periods of both high and low interest rates consistent with what we have observed in the past;

5. suitably parametrised, par yields for long-dated bonds should have a reasonable probability of attaining both high and low values consistent with what we have observed in the past;

6. the model are preferrably time homogeneous (but need not be) and, with an appropriate number of factors, the constant parameters in the model should not need regular recalibration.

Characteristics (4) and (5) depend critically on the values assigned to the mean-reversion rates for the Ornstein-Uhlenbeck processes, $\alpha_1, \ldots, \alpha_n$. In particular, at least one of the α_i must be very low. For example, if α_1 is very low then this means first that the state variable $X_1(t)$ is subject to long cycles. This feeds through to long cycles in rates of interest. In addition the lowest value of the α_i is the one that allows par yields on long-dated bonds to vary over a wide range.

Additional characteristics of the model are:

- the model can produce yield curves similar to those observed in Japan during the early part of the 2000's;

- interest-rate volatility is (very) approximately proportional to the interest rates themselves (for example, the volatility of the short rate, $r(t)$ is roughly proportional to $r(t)$).[2]

As part of our analysis of these models we investigated to what extent volatility might be stochastic. Historical data clearly indicate that short-term volatility in interest rates cannot be explained by the current level of interest rates alone, even though this is a significant factor. Instead volatility also depends on an additional stochastic term $\sigma(t)$. For example, in log-normal models volatility in the short-rate, $r(t)$, may be of the form $\sigma(t)r(t)$.

Evidence for this is presented in Figure 1.1. Here we took 3-month US Treasury Bill data, $y(t)$ (daily observations running from 1953 through to the end of 2001, available online at www.federalreserve.gov). We fitted the one-factor model for $y(t)$ with SDE $dy(t) = -\alpha(y(t) - \mu)dt + \sigma y(t)^\gamma dW(t)$. In contrast to the investigation of this model by Chan *et al.* (1992) we used maximum likelihood estimation. Standardised residuals based on $\hat{\gamma} = 0.7699$ are shown in the lower plot in Figure 1.1. Because the time step is one trading day, these residuals have very little dependence on the linear form of the drift. In this lower plot we can note that to the eye there is a significant

---

[2]It follows that the family of models share certain characteristics with the market models of Brace, Gatarek & Musiela (1997) and Jamshidian (1997).

3-Month Treasury Bill Yields, Daily 1953-2001

Daily Standardised Residuals
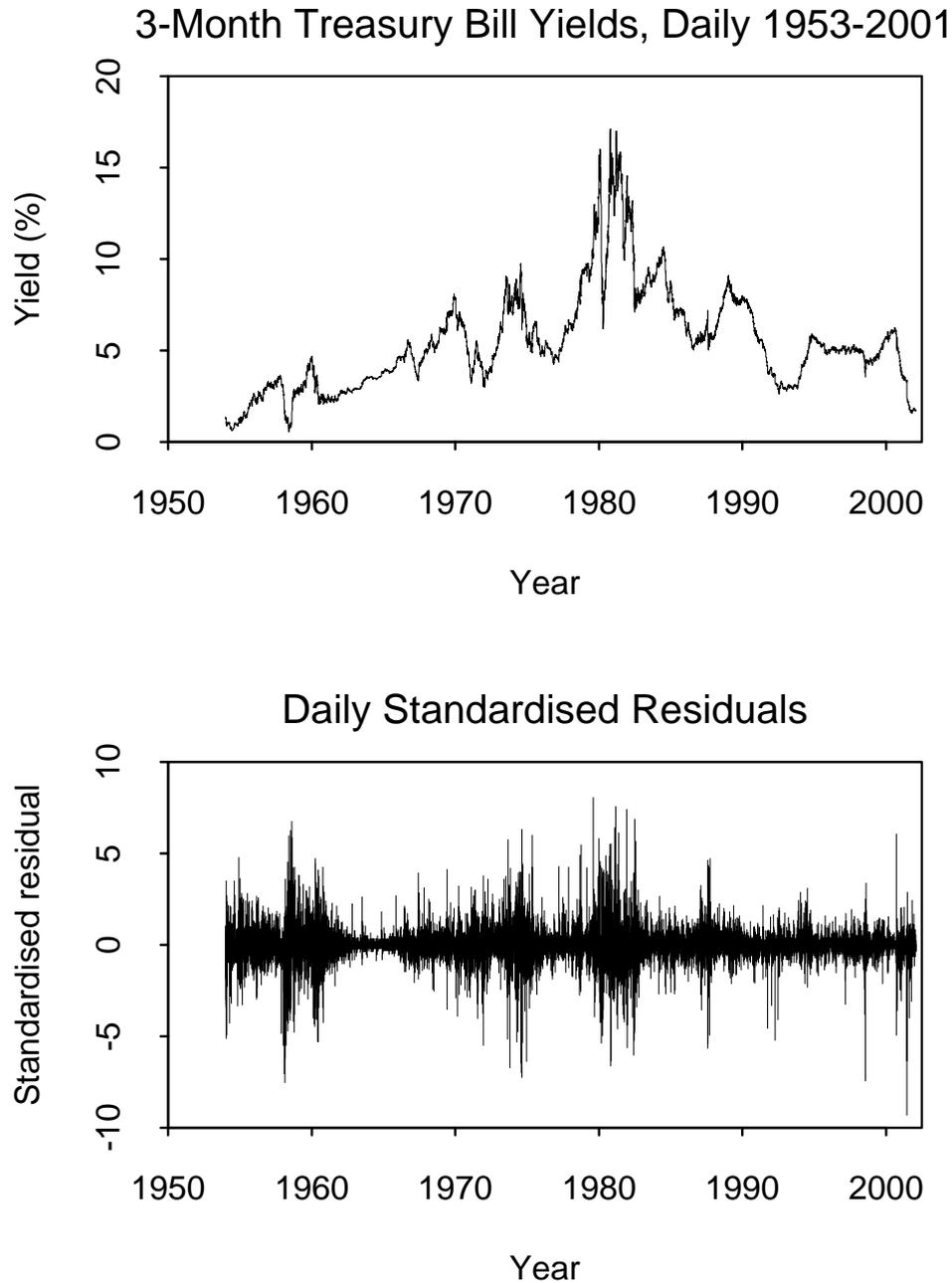
Figure 1.1: US interest rates 1953-2002. Top: daily 3-month Treasury Bill yields. Bottom: daily standardised residuals. Time unit = 1 day. $\epsilon(t) = (y(t) - \hat{y}(t))/\hat{\sigma} y(t-1)^{\hat{\gamma}}$ where $\hat{y}(t) = \hat{\mu} + e^{-\hat{\alpha}}(y(t) - \hat{\mu})$, and $\hat{\mu} = 5.60\%$, $\hat{\alpha} = 0.0003870$ per day, $\hat{\sigma} = 0.02243$ per day and $\gamma = 0.7699$ are the maximum likelihood estimators. Null hypothesis assumes $dy(t) = -\alpha(y(t) - \mu)dt + \sigma y(t)^{\gamma} dW(t)$. Standardised residuals exhibit clear clustering of high and low values indicating the presence of time-varying volatility.

amount of clustering of both high and low residuals. This level of clustering would not seen if the residuals were genuinely indendent and identically distributed.

For the family of models given in equations (1.1) and (1.2) we found that a small amount of stochastic volatility does arise in multifactor models. However, this is rather less than the level of stochastic volatility we see in Figure 1.1. This deficiency was a key driver behind the present work.

In his concluding section, Cairns (2004a) noted that the family of models could be extended to a wider *Integrated Affine* (IA) class. We will see later in this paper how this wider class on models can incorporate a greater degree of stochastic volatility. In the family of models described in equations (1.1) and (1.2) prices depend on an n-dimensional Ornstein-Uhlenbeck process, $X(t)$, through the integral of a log-affine function $H(u, x) = \exp(B(u) + C(u)'X(t))$. Cairns posed the question: are there more general dynamics for $X(t)$ (that is, not Ornstein-Uhlenbeck) which still retain this integral log-affine structure?

Bearing this question in mind, we propose the following time-homogeneous model for zero-coupon-bond prices: for $t < T$,

$$P(t, T, X(t)) = \frac{I_0(t, T, X(t))}{I_0(t, t, X(t))} \tag{1.3}$$

$$\text{where} \quad I_0(t, T, X(t)) = \int_T^\infty M(t, s, X(t)) ds \tag{1.4}$$

$$\text{and} \quad M(t, s, x) = e^{-\beta t + A(s-t) + B(s-t)^T X(t)} \tag{1.5}$$

for some n-dimensional diffusion $X(t)$ and deterministic functions $A(u)$ and $B(u)$. We are interested in discovering for what processes $X(t)$ and forms for $A(u)$ and $B(u)$ is the proposed model for $P(t, T, X(t))$ arbitrage free? Flesaker & Hughston (1996) demonstrated that the model will be arbitrage free if for all T, $M(t, T, X(t))$ is a martingale under the pricing measure, $\hat{P}$, for all $0 \le t \le T$. This is the key to our present work.

We have already seen that this model is arbitrage free when $X(t)$ is an n-dimensional Ornstein-Uhlenbeck process and $A(u)$ and $B(u)$ are appropriate deterministic functions. Additionally, in the case where $n = 1$, Brody & Hughston (2001, 2002) demonstrated that the model is arbitrage free if $X(t)$ has the stochastic differential equation $dX(t) = \alpha(\mu - X(t))dt + \sigma\sqrt{X(t)}d\hat{Z}(t)$: that is, a Cox-Ingersoll-Ross (1985) type of model for $X(t)$.

2 General theory

In this section we will discuss which models fit into the Integrated Affine class. The natural starting point in such an analysis is the work of Duffie & Kan (1996). They

looked at the broad class of multifactor affine term-structure models (including the one-factor models of Vasicek, 1977, and Cox, Ingersoll & Ross, 1985) and provided necessary and sufficient conditions for an affine representation to exist. Here we find that Duffie & Kan's result can be adapted in a simple way to show that the same class of affine models for $X(t)$ are necessary and sufficient for the integrated-affine term-structure models we consider in the present paper.

Theorem 2.1

Let $X(t)$ be some n-dimensional diffusion process giverned by the stochastic differential equation $dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))d\hat{W}(t)$ where $\mu(t, X(t))$ is an $n \times 1$ previsible vector process and $\sigma(t, X(t))$ is an $n \times n$ previsible matrix process.

(a) Suppose that for all T the diffusion processes $M(t, T, X(t)) = \exp[A(t, T) + B(t, T)^T X(t)]$ are strictly positive martingales for all $0 < t < T$, for some scalar deterministic function $A(t, T)$ and $n \times 1$ deterministic function $B(t, T)$.

Then $\mu(t, x)$ and $V(t, x) = \sigma(t, X(t))\sigma(t, X(t))^T$ must be linear (that is, affine) functions of x.

(b) Suppose that $\mu(t, x)$ and $V(t, x) = \sigma(t, X(t))\sigma(t, X(t))^T$ are linear functions of x.

Then there exist deterministic functions $A(t, T)$ and $B(t, T)$ such that for all T the processes $M(t, T, X(t)) = \exp[A(t, T) + B(t, T)^T X(t)]$ are $\hat{P}$-martingales.

Proof: A sketch of the proof which adapts that of Duffie & Kan (1996) is given in the Appendix.

As in Duffie & Kan we need to pay further attention to the structure of $V(t, X(t))$ to ensure that $X(t)$ has a strong solution for all t. In particular, we must ensure that the $v_{ii}(t, X(t))$ remain strictly positive for all t almost surely. Since we are restricted to the same class of affine models as in Duffie & Kan we can quote from their original article the conditions for this to be true.

We will restrict ourselves to the time-homogeneous model for simplicity.

Remark 2.2 *(Duffie & Kan (1996))*

  (a)   (i) The drift of the process $X(t)$ is given by the vector $\mu(x) = \mu_0 + \mu_1 x$ for some $n \times 1$ vector μ_0 and $n \times n$ matrix μ_1.

  (ii) $\sigma(x)$ can be written in the form

$$\sigma(x) = S \begin{pmatrix} \sqrt{w_1(x)} & 0 & \cdots & 0 \\ 0 & \sqrt{w_2(x)} & \cdots & 0 \\ & & \ddots & \\ 0 & \cdots & 0 & \sqrt{w_n(x)} \end{pmatrix}$$

where S is a constant $n \times n$ matrix and the $w_i(x) = \alpha_i + \beta_i^T x$ for scalars $\alpha_1, \ldots, \alpha_n$ and $n \times 1$ vectors $\beta_1, \ldots, \beta_n$. (The non-zero β_i give rise to stochastic volatility.)

(b) For all i:

  (i) For all x such that $w_i(x) = 0$, $\beta_i^T(\mu_0 + \mu_1 x) > \beta_i^T SS^T \beta_i$. (This ensures that boundaries where individual volatilities become equal to zero are never in fact hit: that is, it ensures that all volatilities remain strictly positive for all time. This is the multivariate generalisation of the conditions on the parameters of the Cox-Ingersoll-Ross (1985) model for the short-rate to remain strictly positive.)

  (ii) For all j, if $(\beta_i^T S)_j \neq 0$ then $v_i(x) = k_{ij}v_j(x)$ for some positive constant k_{ij} for all x. (This condition ensures that instantaneous variances when equal to zero cannot be driven negative by dependence on other factors which have non-zero volatilities.)

3 Investigation of a two-factor model

In this section we will investigate the time-homogeneous model

$$P(t,T) = \frac{I_0(t,T,X(t))}{I_0(t,t,X(t))} = \frac{\int_T^\infty M(t,u,X(t))du}{\int_t^\infty M(t,u,X(t))du}$$

$$\text{where } M(t,u) = e^{-\beta t + A(u-t) + B_1(u-t)X_1(t) + B_2(u-t)X_2(t)}$$

$$= e^{-\beta t}H(u-t,X(t)).$$

The state variables $X_1(t)$ and $X_2(t)$ are governed by the SDE's

$$dX_1(t) = \alpha_1(\mu_1 - X_1(t))dt + \sigma_1\sqrt{X_1(t)}d\hat{W}_1(t)$$

$$dX_2(t) = -\alpha_2 X_2(t)dt + \sqrt{X_1(t)}\left(\rho d\hat{W}_1(t) + \sqrt{1-\rho^2}d\hat{W}_2(t)\right)$$

where $\hat{W}_1(t)$ and $\hat{W}_2(t)$ are independent, standard Brownian motions under the pricing measure $\hat{P}$.

For the model to be arbitrage free we require that the family of diffusions $M(t,u,X(t))$ be martingales under $\hat{P}$. By application of Ito's formula we have the SDE for $M(t,u,X(t))$

$$dM(t,u,X(t)) =$$

$$e^{-\beta t}H(u-t,X(t))\Big[\Big(-\beta - A'(u-t) - B_1'(u-t)X_1(t) - B_2'(u-t)X_2(t)\Big)dt$$

$$B_1(u-t)dX_1(t) + B_2(u-t)dX_2(t) + \frac{1}{2}B_1(u-t)^2 d\langle X_1\rangle(t)$$

$$+\frac{1}{2}B_2(u-t)^2 d\langle X_2\rangle(t) + B_1(u-t)B_2(u-t)d\langle X_1, X_2\rangle(t)\Big].$$

If we substitute the known expressions for $dX_1(t)$ etc. and with some rearrangement we get

$$dM(t, u, X(t)) = e^{-\beta t} H(u - t, X(t)) \Big[g_0(u - t, X(t))dt$$

$$+ g_1(u - t, X(t))d\hat{Z}_1(t) + g_2(u - t, X(t))d\hat{Z}_2(t) \Big] \quad (3.1)$$

$$\text{where} \quad g_1(s, x) = B_1(s)\sigma_1\sqrt{x_1} + \rho B_2(s)\sqrt{x_1} \quad (3.2)$$

$$g_2(s, x) = B_2(s)\sqrt{1 - \rho^2}\sqrt{x_1} \quad (3.3)$$

$$\text{and} \quad g_0(s, x) = -\beta - A'(s) - x_1 B_1'(s) - x_2 B_2'(s) \quad (3.4)$$

$$+ B_1(s)\alpha_1(\mu_1 - x_1) - B_2(s)\alpha_2 x_2$$

$$+ \frac{1}{2}B_1(s)^2\sigma_1^2 x_1 + \frac{1}{2}B_2(s)^2 x_1 + B_1(s)B_2(s)\rho\sigma_1 x_1.$$

For $M(t, u, X(t))$ to be a martingale it is necessary that $g_0(s, x) = 0$ for all s and x. Since $g_0(s, x)$ is linear in x_1 and x_2 this means that

$$-\beta - A'(s) + \alpha_1\mu_1 B_1(s) = 0 \quad (3.5)$$

$$-B_1'(s) - \alpha_1 B_1(s) + \frac{1}{2}\sigma_1^2 B_1(s)^2 + \frac{1}{2}B_2(s)^2 + \rho\sigma_1 B_1(s)B_2(s) = 0 \quad (3.6)$$

$$\text{and} \quad -B_2'(s) - \alpha_2 B_2(s) = 0. \quad (3.7)$$

From (3.7) we have the solution

$$B_2(s) = B_2(0)e^{-\alpha_2 s}.$$

From (3.6) we get

$$B_1'(s) = -\alpha_1 B_1(s) + \frac{1}{2}\sigma_1^2 B_1(s)^2 + \frac{1}{2}B_2(s)^2 + \rho\sigma_1 B_1(s)B_2(s). \quad (3.8)$$

Because of the form of $B_2(s)$ this requires numerical solution (which is computationally straightforward and fast). Similarly (3.5) gives us

$$A'(s) = -\beta + \alpha_1\mu_1 B_1(s).$$

Without loss of generality we may assume that $A(0) = 0.$[3] The initial conditions for $B_1(0)$ and $B_2(0)$ do have a material impact on the model and can therefore be treated as model parameters.

Asymptotically we can note (by inspection of equation 3.8) that $B_1(s) = O(e^{-\alpha_1 s})$ if $\alpha_1 < 2\alpha_2$ (since the $-\alpha_1 B_1(s)$ term dominates and $B_1(s) = O(e^{-2\alpha_2 s})$ if $\alpha_1 > 2\alpha_2$ since the $\frac{1}{2}B_2(u)$ term dominates. If $\alpha_1 > 2\alpha_2$ then numerical experiments suggest that, given $B_2(0)$ the limiting value of $B_1(s)e^{2\alpha_2 s}$ is independent of the value of $B_1(0)$. This is not true when $\alpha_1 < 2\alpha_2$. We can also see that $A(s) = -\beta s + \tilde{A}(s)$ where $\tilde{A}(s)$ tends to some constant as s tends to infinity.

---

[3]A different value for $A(0)$ would multiply all of the $M(t, u, X(t))$ by a constant factor which then gets cancelled out when we take the ratios of the integrals to calculate prices. That is, different values of $A(0)$ have no impact on the term structure.

3.1   Volatility term structure

Following on from Flesaker & Hughston (1996) and Cairns (2004a) we can derive bond and interest-rate volatilities as follows. Since $M(t, u, X(t))$ is a strictly positive martingale there exists a previsible volatility process for $M(t, u, X(t))$ such that

$$dM(t, u, X(t)) = M(t, u, X(t))\sigma_M(t, u)d\hat{W}(t).$$

Now define the vector process

$$V(t, T) = V(t, T, X(t)) = \frac{\int_T^\infty M(t, u, X(t))\sigma_M(t, u)du}{\int_T^\infty M(t, u, X(t))du}.$$

Then the volatility of the zero-coupon bond $P(t, T)$ is

$$S_P(t, T) = V(t, T) - V(t, t).$$

For instantaneous forward rates, $f(t, T)$, the SDE is

$$df(t, T) \;\;=\;\; f(t, T)\left(\sigma_M(t, T) - V(t, T)\right)'\left\{d\hat{W}(t) - V(t, T)dt\right\}.$$

This implies that the forward-rate volatility is

$$\sigma_f(t, T) \;\;=\;\; f(t, T)(\sigma_M(t, T) - V(t, T)). \qquad (3.9)$$

For a given t and T the individual components of the vector $\sigma_f(t, T)$ are independent volatilities. These equations hold for general volatility functions $\sigma_M(t, T)$. In the case of the two-factor model being presented here we can refer to equations (3.1) to (3.3) to see that

$$
\begin{aligned}
\sigma_{M1}(t, T) &= s_{M1}(T - t)\sqrt{X_1(t)} \\
\sigma_{M2}(t, T) &= s_{M2}(T - t)\sqrt{X_1(t)} \\
\text{where } s_{M1}(T - t) &= B_1(T - t)\sigma_1 + \rho B_2(T - t) \\
\text{and } s_{M2}(T - t) &= \sqrt{1 - \rho^2}B_2(T - t)
\end{aligned}
$$

are deterministic functions. Hence $V(t, T) = (V_1(t, T), V_2(t, T))^T$ where

$$
\begin{aligned}
V_j(t, T) &= \frac{I_{1j}(t, T, X(t))\sqrt{X_1(t)}}{I_0(t, T, X(t))} \\
I_0(t, T, X(t)) &= \int_T^\infty M(t, u, X(t))du = e^{-\beta t}\int_{T-t}^\infty H(v, X(t))dv \\
\text{and} \\
I_{1j}(t, T, X(t)) &= \int_T^\infty M(t, u, X(t))s_{Mj}(u - t)du = e^{-\beta t}\int_{T-t}^\infty H(v, X(t))s_{Mj}(v)dv.
\end{aligned}
$$

Hence we arrive at the forward-rate volatility functions

$$\sigma_{fj}(t,T) = f(t,T)s_{fj}(T-t)\sqrt{X_1(t)}$$

$$\text{where } s_{fj}(T-t) = \left(s_{Mj}(T-t) - \frac{I_{1j}(t,T,X(t))}{I_0(t,T,X(t))}\right).$$

We will refer to the $s_{fj}(u)$ as the independent log-volatility functions.

From these equations we can note the following points.

- The principle stochastic element in the forward-rate volatilities is $\sqrt{X_1(t)}$. There is an additional contribution from $I_{1j}(t,T,X(t))/I_0(t,T,X(t))$ but this is much less significant.[4]

- Given $X_1(t)$ the volatility of $f(t,T)$ is approximately proportional to $f(t,T)$ itself. There is, again, an additional contribution from $I_{1j}(t,T,X(t))/I_0(t,T,X(t))$ but this is much less significant.

- The main determinants of the term-structure of volatility are the deterministic functions $s_{mj}(T-t)$.

From these observations we can conjecture that the model shares important characteristics with log-normal interest-rate models and market models with stochastic volatility (see, for example, Rebonato, 2002, and Joshi & Rebonato, 2003).

3.2 Numerical investigations

We will now illustrate various characteristics of the two-factor model. We will use as our central parameter set the values $\mu_1 = 1$, $\alpha_1 = 2$, $\sigma_1 = 0.5$, $\alpha_2 = 0.2$, $\rho = 0.2$, $\beta = 0.05$, $B_1(0) = 0.5$ and $B_2(0) = 0.2$. (Recall that $B_1(0)$ and $B_2(0)$ do have a material impact on the results and are therefore just as much parameters in the model as, say, α_1.)

In Figure 3.1 we plot sample spot rate curves $R(0,t) = -t^{-1}\log P(0,t)$ for a range of values for $X(t)$. In the upper plot (A) we see how $X_1(0)$ affects spot rates. The range enclosed by the dotted curve gives a 95% unconditional confidence interval given $X_2(0) = 0$.[5] Thus we see that its value influences principally short-term rates and, even then, only in a limited way. The shaded region indicates values for $R(0,t)$ which cannot be attained for the given value of $X_2(0) = 0$, with the lower attainable boundary arising when $X_1(0) = 0$. (Recall that $X_1(t)$ cannot become negative.)

---

[4]For the parameter sets we considered we found that the absolute value of the $I_{1j}(t,T,X(t))/I_0(t,T,X(t))$ was relatively small compared with the $s_{Mj}(T-t)$. Thus, although the $I_{1j}(t,T,X(t))/I_0(t,T,X(t))$ do vary with both $X_1(t)$ and $X_2(t)$, the impact of this variation on $s_{Mj}(T-t) - I_{1j}(t,T,X(t))/I_0(t,T,X(t))$ is small.

[5]Since $\rho \neq 0$ this is not strictly an unconditional confidence interval. However, it gives a good indication of the range of values taken by $X_1(t)$ 95% of the time.
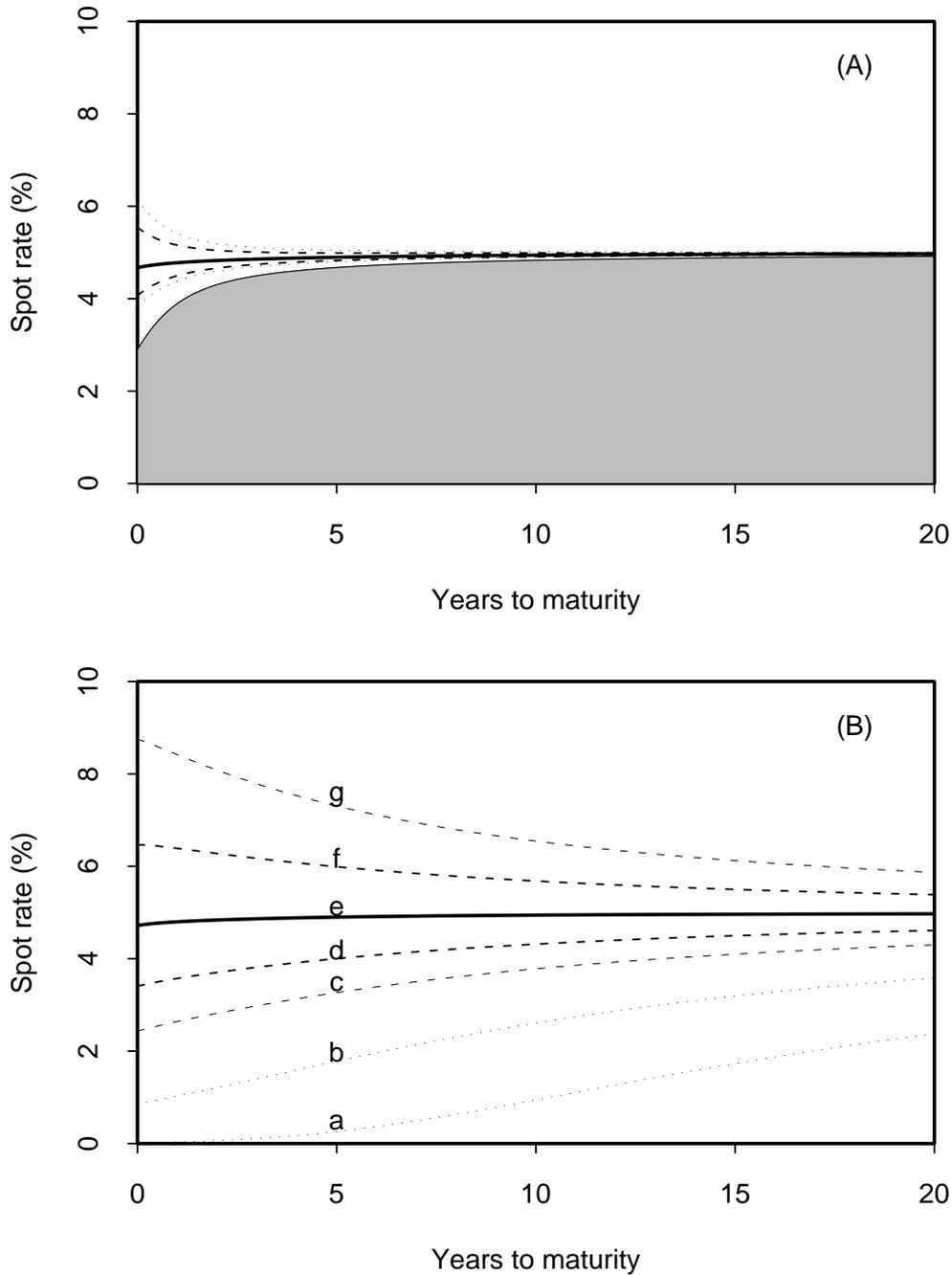
Figure 3.1: Sample spot rate curves for the two-factor model with stochastic volatility. (A) Given $X_2(0) = 0$, spot rate curves when (bottom to top) $X_1(0) =$0, 0.572, 0.696, 0.979, 1.331, and 1.546 (corresponding to the 0%, 2.5%, 10%, 50%, 90% and 97.5% quantiles of the stationary distribution for $X_1(t)$). The shaded region shows the unattainable interest rates for the given value of $X_2(0)$. (B) Given $X_1(0) = 1$, spot rate curves for $X_2(0) =$-30, -10, -4, -2, 0, 2, and 4 (curves a to g respectively). Parameters: $\mu_1 = 1$, $\alpha_1 = 2$, $\sigma_1 = 0.5$, $\alpha_2 = 0.2$, $\rho = 0.2$, $\beta = 0.05$, $B_1(0) = 0.5$ and $B_2(0) = 0.2$.
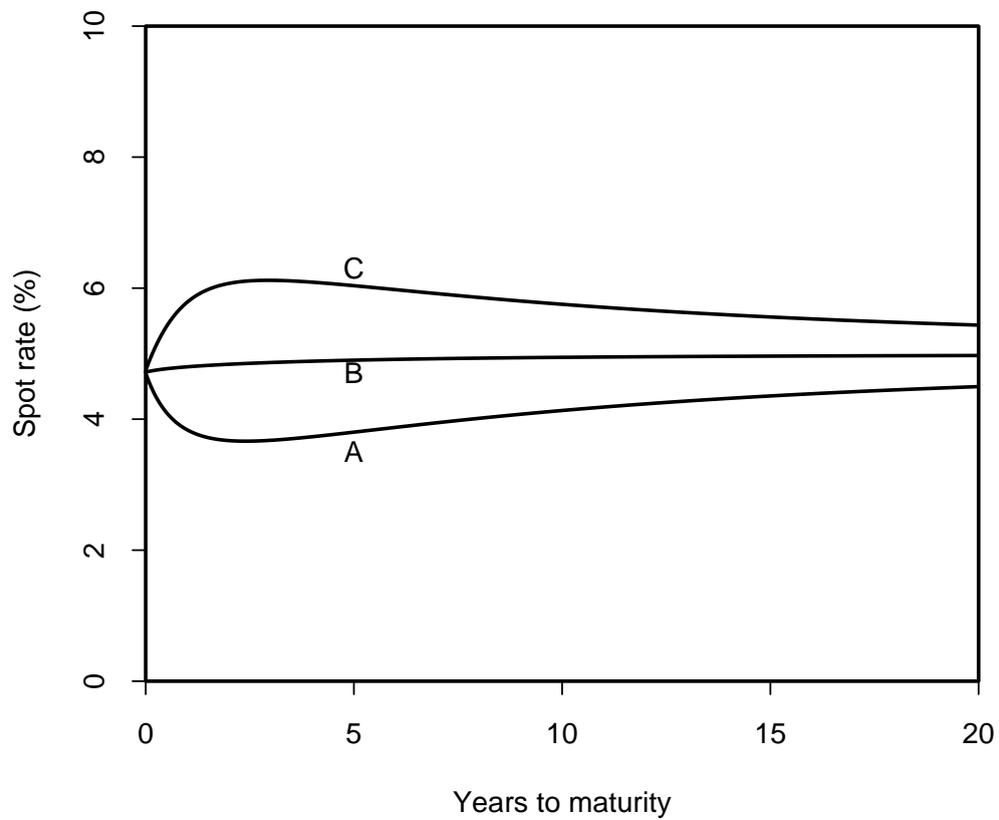
Figure 3.2: Sample spot rate curves for the two-factor model with stochastic volatility. $(X_1(0), X_2(0)) = $ A: $(2, -1)$, B: $(1, 0)$, C: $(0.12.5)$. Parameters: $\mu_1 = 1$, $\alpha_1 = 2$, $\sigma_1 = 0.5$, $\alpha_2 = 0.2$, $\rho = 0.2$, $\beta = 0.05$, $B_1(0) = 0.5$ and $B_2(0) = 0.2$.

In the lower plot (B) in Figure 3.1 we see how the spot rates depend on $X_2(0)$ when $X_1(0) = 1$. Curves (c), (d), (e), (f) and (g) represent likely outcomes. Curves (b) and (a) are more extreme cases. They are extremely unlikely but nevertheless are possible. Curve (a), for example, is similar to the current yield curve in Japan. As $X_2(0)$ becomes more and more negative the spot rate curve continues to flatten out indicating that spot rates for all maturities can get arbitrarily close to zero. This sort of limiting behaviour is similar to, for example, the Black & Karasinski (1991) model but different from, for example, the Cox-Ingersoll-Ross (1985) model (where only the risk-free rate can attain values arbitrarily close to 0).

Figure 3.1 shows is that we can get typical rising, level and falling spot-rate curves. For different combinations (see Figure 3.2) of $X_1(0)$ and $X_2(0)$ we can also produce humped and dipped curves. However, we should stress that the current two-factor model is used here for illustrative purposes. The shape of the yield curve is primarily dependent on the single factor $X_2(0)$. If we feel that a richer variety of yield curves is required (which is the case in most applications) then it is more appropriate to use, say, a third factor, $X_3(t)$, rather than rely on $X_1(t)$ to manipulate the shape. The latter approach is likely to distort the role of $X_1(t)$ as the primary determinant of interest-rate volatility.

In Figure 3.3 we show contour plots for different interest rates as a means of investigating how these rates depend on the values of $X_1(t)$ and $X_2(t)$. For the risk-free rate of interest, $r(t)$ (top left), we can make two observations. First, there is significant dependence on both $X_1(t)$ and $X_2(t)$. Second, the absolute range of values which are likely to arise is reasonably wide. Also plotted are 10,000 simulated values of $(X_1(10), X_2(10))'$ given $(X_1(0), X_2(0))' = (1,0)'$ to allow us to judge which points are likely and which are not. There is a small positive correlation reflecting the short term correlation ($\rho = 0.2$) between $X_1(t)$ and $X_2(t)$.

In contrast to $r(t)$ we see that the 10-year forward rate (top right) varies within a very much smaller range and that there is almost no dependence on $X_1(t)$.

In between these extremes we have the 10-year spot rate and the 30-year par yield. Each varies over a medium range, has a small amount of dependence on $X_1(t)$ and depends mainly on $X_2(t)$.

In Figure 3.4 we plot independent log-volatilities for the forward-rate curve. In the top plot we show the volatility term structure when $X(0) = (1,0)'$ and add to this the total log-volatility curve $s_f(T) = \sqrt{s_{f1}(T)^2 + s_{f2}(T)^2}$. From this we can see that volatility of short-term forward rates is influenced most by $d\hat{Z}_1(t)$. The impact of this term declines rapidly as t increases. In contrast $s_{f2}(T)$ declines much less rapidly and $d\hat{Z}_2(t)$ is the dominant influence for medium and long-term maturities. These observations are consistent with our comments on Figure 3.3. The short-term impact of $X_1(t)$, though, tends to be proportionately lower in Figure 3.3 because the high value of α_1 means that the impact of shocks $d\hat{Z}_1(t)$ persist for a much shorter period than the $d\hat{Z}_2(t)$.
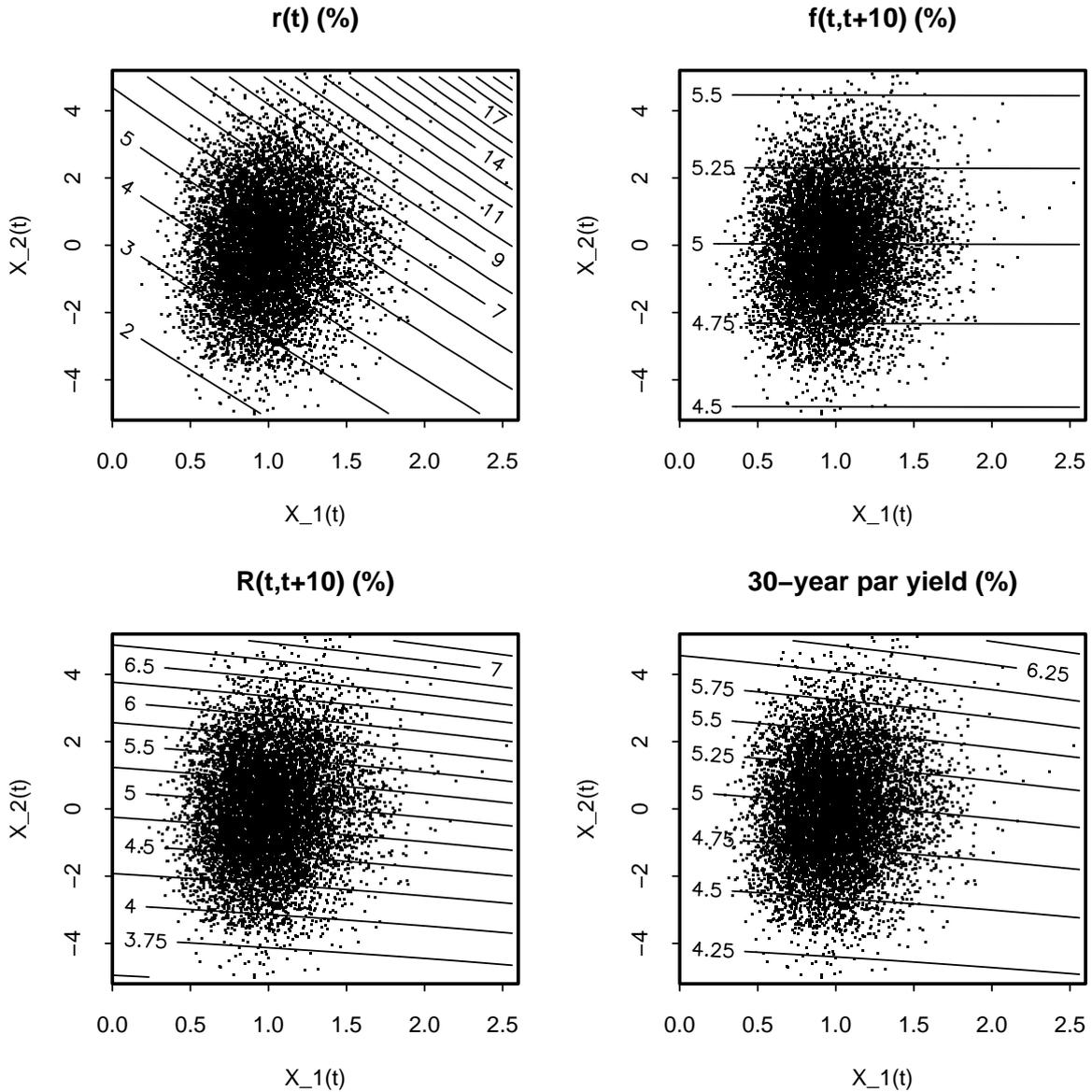
Figure 3.3: Dependence of different interest rates on $X_1(t)$ and $X_2(t)$. Each plot shows contours connecting pairs of values for $(X_1(t), X_2(t))^T$ which result in the same rate of interest. Each plot also includes a scatter plot of 10000 simulated values of $(X_1(10), X_2(10))^T$ given $(X_1(0), X_2(0))^T = (1, 0)^T$. Top left: risk-free rate $r(t)$. Top right: 10-year forward rate $f(t, t + 10)$. Bottom left: 10-year spot rate $R(t, t + 10)$. Bottom right: 30-year par yield $\rho(t, t + 30)$. Parameters: $\mu_1 = 1$, $\alpha_1 = 2$, $\sigma_1 = 0.5$, $\alpha_2 = 0.2$, $\rho = 0.2$, $\beta = 0.05$, $B_1(0) = 0.5$ and $B_2(0) = 0.2$.
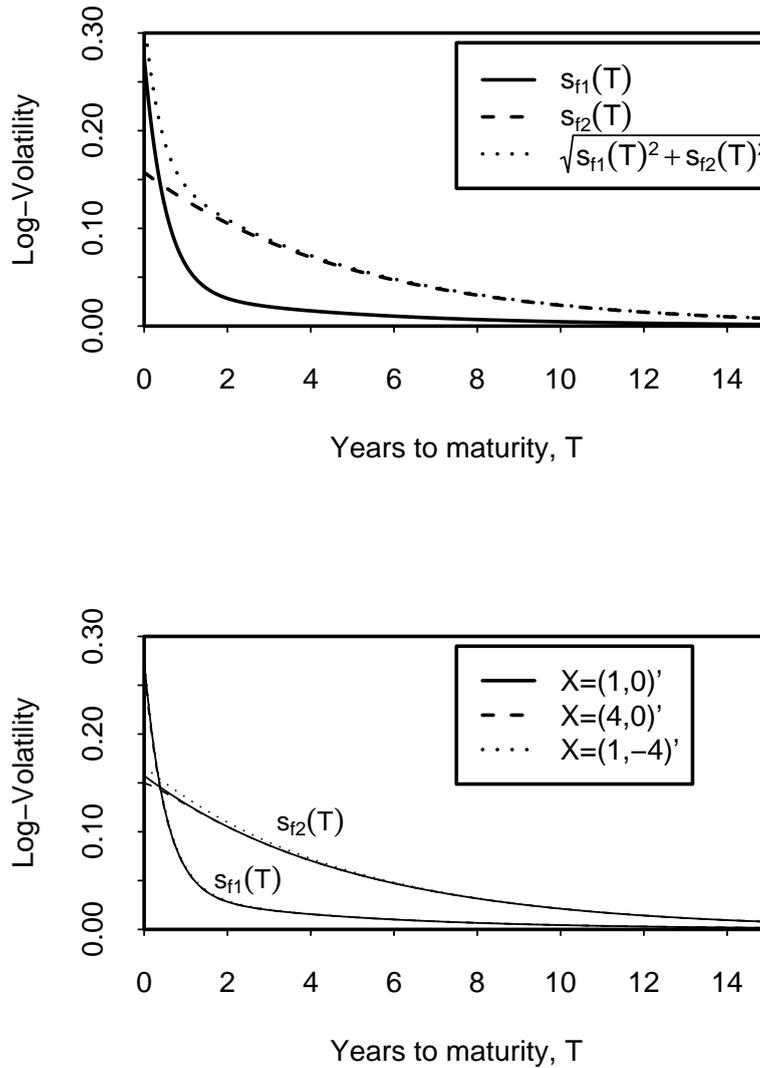
Figure 3.4: Forward-rate volatility term structure. Independent log-volatility functions $s_{f1}(T)$ and $s_{f2}(T)$ and the total forward-rate log-volatility function $s_f(T) = \sqrt{s_{f1}(T)^2 + s_{f2}(T)^2}$. Top: forward rate volatilities when $X(0) = (1,0)^T$. Bottom: impact on independent log-volatilities of different values for $X_1(0)$ and $X_2(0)$. Parameters: $\mu_1 = 1$, $\alpha_1 = 2$, $\sigma_1 = 0.5$, $\alpha_2 = 0.2$, $\rho = 0.2$, $\beta = 0.05$, $B_1(0) = 0.5$ and $B_2(0) = 0.2$.

In the lower plot in Figure 3.4 we show how changes in the values of $X_1(0)$ and $X_2(0)$ affect the independent log-volatility functions $s_{f1}(T)$ and $s_{f2}(T)$. The dashed curves show what happens in $X_1(0)$ is increased substantially. The impact of this change is only visible for $s_{f2}(T)$ at very short maturities and not at all for $s_{f1}(T)$. The impact of changing $X_2(0)$ from 0 to -4 (dotted curves) is also relatively small, but it is more visible. We can conclude then that the deterministic components of the forward-rate volatilities dominate the independent log-volatility term-structure.

4   A three-factor model

We will now describe briefly a three-factor model with one source of stochastic volatility. Specifically we will show how this enriches the range of outcomes relative to the two-factor model.

The dynamics of the state variables are

$$dX_1(t) = \alpha_1(\mu_1 - X_1(t))dt + \sigma_1\sqrt{X_1(t)}\sum_{j=1}^{3}c_{1j}d\hat{W}_j(t)$$

$$dX_2(t) = -\alpha_2 X_2(t)dt + \sqrt{X_1(t)}\sum_{j=1}^{3}c_{2j}d\hat{W}_j(t)$$

$$dX_3(t) = -\alpha_3 X_3(t)dt + \sqrt{X_1(t)}\sum_{j=1}^{3}c_{3j}d\hat{W}_j(t)$$

where, for each i, $\sum_{j=1}^{3}c_{ij}^2 = 1$. $\hat{W}(t)$ is a standard 3-dimensional Brownian motion under the pricing measure $\hat{P}$. If we define the matrix $C = (c_{ij})_{i,j=1}^{3}$ then $CC' = \rho = (\rho_{ij})_{i,j=1}^{3}$ is the instantaneous correlation matrix, between the state variables.

Without loss of generality we may assume that C is lower triangular (since it is the instantaneous correlation matrix ρ that really matters).

As before we define

$$M(t,T,X(t)) = e^{-\beta t}H(T-t,X(t)$$
$$\text{where } H(u,x) = e^{A(u)+B_1(u)X_1(t)+B_2(u)X_2(t)+B_3(u)X_3(t)}.$$

If we follow the same arguments as before then the $M(t,T,X(t))$ are martingales

under $\hat{P}$ if

$$B_2'(u) \;=\; -\alpha_2 B_2(u) \tag{4.1}$$

$$B_3'(u) \;=\; -\alpha_3 B_3(u) \tag{4.2}$$

$$B_1'(u) \;=\; -\alpha_1 B_1(u) + \frac{1}{2}\sigma_1^2 B_1(u)^2 + \frac{1}{2}B_2(u)^2 + \frac{1}{2}B_3(u)^2 + \sigma_1\rho_{12}B_1(u)B_2(u)$$
$$\qquad\quad +\sigma_1\rho_{13}B_1(u)B_3(u) + \rho_{23}B_2(u)B_3(u) \tag{4.3}$$

$$A'(u) \;=\; -\beta + \alpha_1\mu_1 B_1(u). \tag{4.4}$$

From (4.1) and (4.2) we get $B_2(u) = B_2(0)e^{-\alpha_2 u}$ and $B_3(u) = B_3(0)e^{-\alpha_3 u}$. Equations (4.3) and (4.4) can be solved numerically. We then have

$$dM(t,T,X(t)) \;=\; M(t,T,X(t))\sqrt{X_1(t)}\sum_{j=1}^{3} s_{Mj}(T-t)^T d\hat{W}(t)$$

$$\text{where }\; s_{M1}(u) \;=\; \sigma_1 B_1(u) + c_{21}B_2(u) + c_{31}B_3(u)$$

$$s_{M2}(u) \;=\; c_{22}B_2(u) + c_{32}B_3(u)$$

$$s_{M3}(u) \;=\; c_{33}B_3(u).$$

Sample spot-rate curves are plotted in Figure 4.1 for the parameter values $\alpha_1 = 2$, $\alpha_2 = 1.5$, $\alpha_3 = 0.075$, $\sigma_1 = 0.3$, $\mu_1 = 1$, $\beta = 0.05$, $\rho_{12} = 0$, $\rho_{13} = 0$, $\rho_{23} = -0.8$, $B_1(0) = 0.01$, $B_2(0) = 0.25$, $B_3(0) = 0.35$. In this figure we can see that $X_2(0)$ has most impact on short-maturity spot rates. $X_3(0)$ also has a big impact on short-maturity rates but, more importantly, it is effectively the sole determinant of long-maturity spot rates. Thes observations are consistent with the high and low values attached to α_2 and α_3 respectively. In contrast, for the chosen parametrisation, $X_1(0)$ has almost no effect on spot rates: the three curves are almost indistinguishable even for the relatively extreme values tested for $X_1(0)$. (The 2.5% and 97.5% unconditional quantiles for $X_1(t)$ are 0.728 and 1.315 respectively.) In effect our choice of parameters means that $X_1(t)$ only affects local volatility in a significant way. Finally we see that we can vary $X_2(0)$ and $X_3(0)$ in ways which can produce both humped and dipped curves as well as level, rising and falling curves.

Forward-rate volatilities can be expressed as in equation (3.9), and based on this we can derive the independent volatilities

$$\sigma_{fj}(t,T) \;=\; f(t,T)s_{fj}(T-t)\sqrt{X_1(t)}\;\text{ for } j = 1,2,3$$

$$\text{where }\; s_{fj}(T-t) \;=\; \left(s_{Mj}(T-t) - \frac{I_{1j}(t,T,X(t))}{I_0(t,T,X(t))}\right)$$

$$\text{and }\; I_{1j}(t,T,X(t)) \;=\; e^{-\beta t}\int_{T-t}^{\infty} H(u,X(t))s_{Mj}(u)du$$

as before.
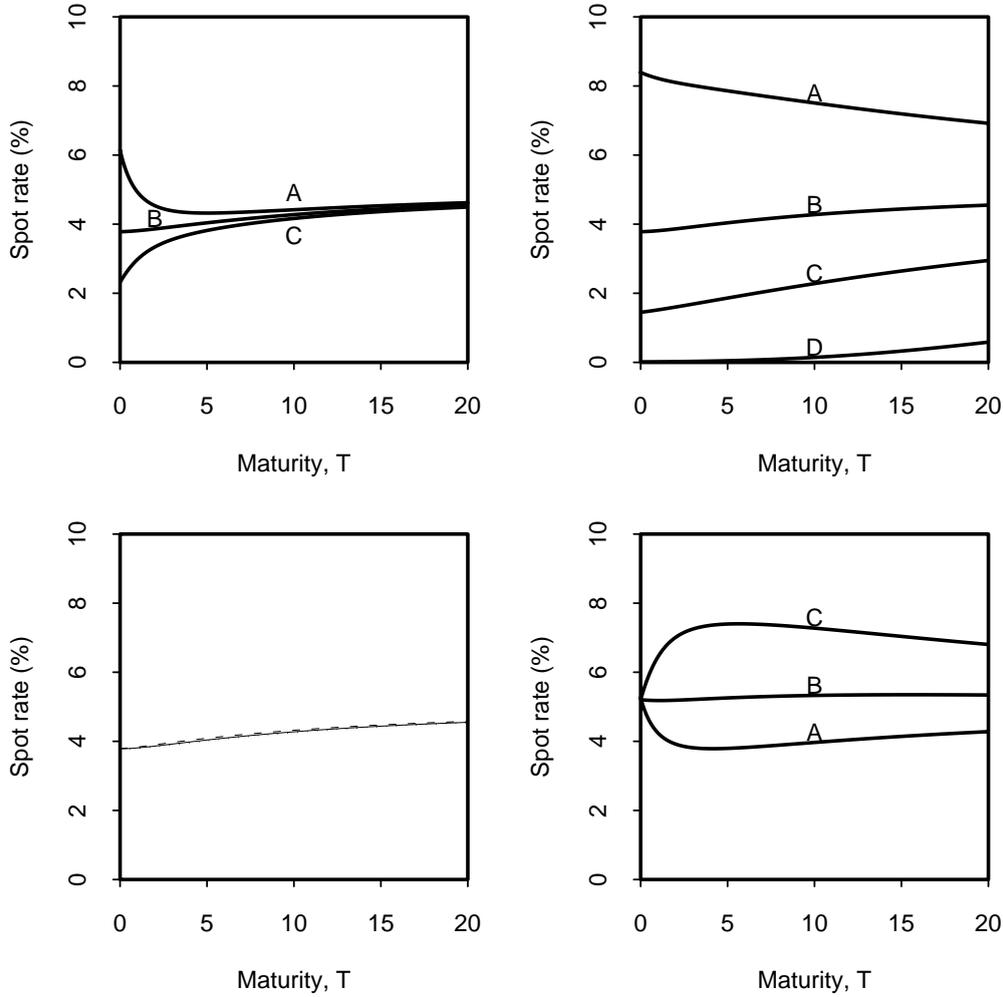
Figure 4.1: Sample spot-rate curves for the three-factor model. Top left: impact of variation in $X_2(0)$. X = A: $(1, 2, 0)$; B: $(1, 0, 0)$; C: $(1, -2, 0)$. Top right: impact of variation in $X_3(0)$. X = A: $(1, 0, 4)$; B: $(1, 0, 0)$; C: $(1, 0, -4)$; D: $(1, 0, -20)$ (extreme case). Bottom left: impact of variation in $X_1(0)$. X = dashed line $(2, 0, 0)$; solid line $(1, 0, 0)$; dotted line $(0.5, 0, 0)$. The three lines are almost indistinguishable. Bottom right: different shapes of curve by varying both $X_2(0)$ and $X_3(0)$. X = A: $(1, 2, -0.7)$; B: $(1, 0, 1.5)$; C: $(1, -2, 4)$. Parameter values: $\alpha_1 = 2$, $\alpha_2 = 1.5$, $\alpha_3 = 0.075$, $\sigma_1 = 0.3$, $\mu_1 = 1$, $\beta = 0.05$, $\rho_{12} = 0$, $\rho_{13} = 0$, $\rho_{23} = -0.8$, $B_1(0) = 0.01$, $B_2(0) = 0.25$, $B_3(0) = 0.35$.
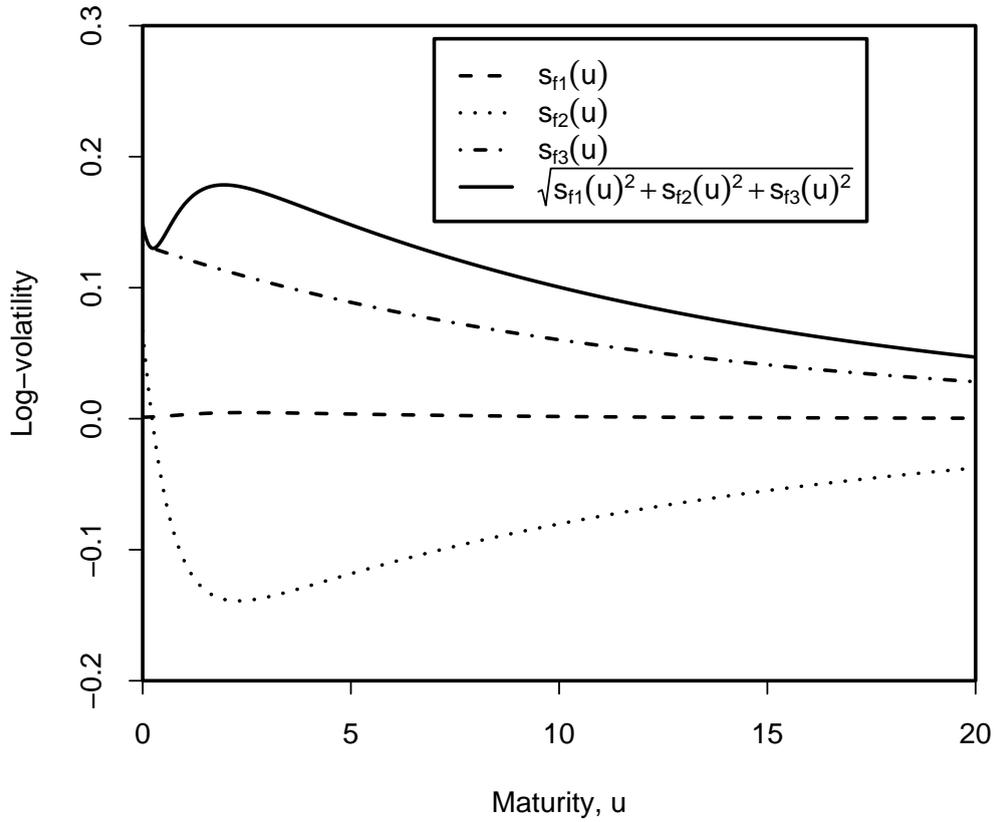
Figure 4.2: Term-structure of forward-rate volatilities for the three-factor model. Independent forward-rate log-volatility curves, $s_{f1}(u)$, $s_{f2}(u)$ and $s_{f3}(u)$ and the total volatility curve $s_f(u)$. Parameter values: $\alpha_1 = 2$, $\alpha_2 = 1.5$, $\alpha_3 = 0.075$, $\sigma_1 = 0.3$, $\mu_1 = 1$, $\beta = 0.05$, $\rho_{12} = 0$, $\rho_{13} = 0$, $\rho_{23} = -0.8$, $B_1(0) = 0.01$, $B_2(0) = 0.25$, $B_3(0) = 0.35$.

In Figure 4.2 we have plotted the independent forward-rate log-volatility curves, $s_{f1}(u)$, $s_{f2}(u)$ and $s_{f3}(u)$ (dashed, dotted and dot-dash curves respectively) in addition to the total log-volatility curve $s_f(u) = \sqrt{s_{f1}(u)^2 + s_{f2}(u)^2 + s_{f3}(u)^2}$ when $X(0) = (1, 0, 0)'$. The shape of the total-log-volatility curve is typical of what we observe in practice, with a peak at around the 2-year maturity mark (see, for example, Brigo & Mercurio, 2001, Rebonato, 2002, and Joshi & Rebonato, 2003). We can see that $s_{f1}(u)$ is generally very low indicating, again, that *changes* in $X_1(t)$ do not affect significantly forward rates. (However, volatility is proportional to $\sqrt{X_1(t)}$, so that it has a significant impact in other ways.) The basic humped shape of the total volatility curve in fact depends on primarily on $B_2(0)$, $B_3(0)$ and ρ_{23}. If $B_2(0) > 0$ and $B_3(0) > 0$ then the humped structure can only occur if $\rho_{23} < 0$. Here we have chosen $B_2(0)$ to produce not just a significant hump but also a small 'hook' at 3-months. This hook is consistent with the emprirical observation that 1 and 2-month LIBOR rates tend to be more volatile than 3-month rates. With different choices for $B_2(0)$, $B_3(0)$ and ρ_{23} we can remove this hook, emphasize the hump or remove the hump entirely.

The shape of $s_{f3}(u)$ indicates that the $d\hat{W}_3(t)$ shocks result in changes in the level and slope of the forward-rate curve. The shape of $s_{f2}(u)$ indicates that the $d\hat{W}_2(t)$ shocks causes humps and dips to increase or decrease.[6]

5 Further comments

We have analysed here two and three-factor models in which one factor gives rise to stochastic volatility. Even with a two-plus-one model as in Section 4 there are limitations which suggest that further factors might be appropriate.

- In the three factor model we chose a high value for α_2 and a low value for α_3 to produce a term-structure for volatility with a hump at around the right maturity. A consequence of α_3 being the only mean-reversion parameter with a low value is that changes in forward rates for maturities over two years are highly correlated.

  This can be mitigated by incorporating a further factor $X_4(t)$ with SDE $dX_4(t) = -\alpha_4 X_4(t)dt + \sum)j = 1^4 d\hat{W}_j(t)$ where α_4 is also quite low.

- In the models analysed the single stochastic volatility $X_1(t)$ means that the *shape* of the term-structure of volatility is broadly fixed. In reality we may find that the shape as well as its overall magnitude evolves over time.

---

[6]However, we can always reparametrise the matrix C while keeping $\rho = CC^T$ as before, in which case the impact of $d\hat{W}_2(t)$ and $d\hat{W}_3(t)$ will be different. The overall impact will, of course, be unchanged.

We can add a stochastic element to the shape of the term structure of volatility by adding a second stochastic volatility process in addition to $X_1(t)$.

A   Proof of Theorem 2.1

(a)

In the equations that follow $A'(t,T)$ represents the partial derivative of A with respect to t *etc.* while $B(t,T)^T$ represents the transpose of the vector B (or the transpose of a matrix). We have

$$
\begin{aligned}
M(t,T,X(t)) &= e^{A(t,T)+B(t,T)^T X(t)} \\
\Rightarrow dM(t,T,X(t)) &= M(t,T,X(t))\Big[\Big(A'(t,T) + B'(t,T)X(t) + B(t,T)^T\mu(t,X(t)) \\
&\qquad + \frac{1}{2}B(t,T)^T\sigma(t,X(t))\sigma(t,X(t))^T B(t,T)\Big)dt \\
&\qquad + B(t,T)\sigma(t,X(t))d\hat{W}(t)\Big]
\end{aligned}
$$

Since $M(t,T,X(t))$ is a $\hat{P}$-martingale we have

$$
\begin{aligned}
a_L(t,T,X(t)) &= a_R(t,T,X(t)) \\
\text{where } a_L(t,T,x) &= -A'(t,T) - B'(t,T)^T x \\
a_R(t,T,x) &= B(t,T)^T\mu(t,x) + \frac{1}{2}B(t,T)^T V(t,x)B(t,T) \\
\text{and } V(t,x) &= \sigma(t,x)\sigma(t,x)^T = \big(v_{ij}(t,x)\big)_{i,j=1}^n.
\end{aligned}
$$

Following Duffie & Kan (1996) we construct a column vector

$$
g(t,x) = (\mu_1(t,x),\ldots,\mu_n(t,x),v_{11}(t,x),\ldots,v_{1n}(t,x),v_{22}(t,x),\ldots,v_{nn}(t,x))^T
$$

where we include only the upper triangular entries in $V(t,x)$. $g(t,x)$ contains $N = n + \frac{1}{2}n(n+1)$ terms. There exists a row vector $h(t,T)$ such that $a_R(t,T,x) = h(t,T)g(t,x)$: for example, $h_1(t,T,x) = B_1(t,T)$, $h_{n+1}(t,T) = \frac{1}{2}B_1(t,T)^2$, and $h_{n+2}(t,T) = B_1(t,T)B_2(t,T)$ *etc.*.

Now $a_L(t,T,x)$ is linear (affine) in $x1,\ldots,x_n$. It follows that $a_R(t,T,x)$ must also be linear in $x1,\ldots,x_n$.

Next consider N maturity dates $T_1 < T_2 < \ldots < T_N < \infty$. Let $h_{ij} = h_j(t,T_i)$ and define the matrix $H = (h_{ij})_{i,j=1}^n$. Then

$$
a_L(t,x) = \begin{pmatrix} a_L(t,T_1,x) \\ \vdots \\ a_L(t,T_N,x) \end{pmatrix} = \begin{pmatrix} a_R(t,T_1,x) \\ \vdots \\ a_R(t,T_N,x) \end{pmatrix} = Hg(x).
$$

Provided H is non-singular this implies that

$$
g(x) = H^{-1}a_L(t,x).
$$

Now H is independent of x, whereas the vector $a_L(t,x)$ is affine in x. Therefore $g(x)$ must be affine in x: that is,

- for each i, $\mu_i(t, x)$ is affine in x;

- for each i and j, $v_{ij}(t, x)$ is affine in x.

(b)

Now suppose that $\mu(t, x)$ and $V(t, x) = \sigma(t, X(t))\sigma(t, X(t))^T$ are linear functions of x. Suppose that $M(t, T, X(t)) = e^{A(t,T)+B(t,T)^T X(t)}$ for some deterministic functions A and B. For $M(t, T, X(t))$ to be a $\hat{P}$-martingale we require $a_L(t, T, X(t)) = a_R(t, T, X(t))$ as in the proof of part (a), to ensure that the drift of $M(t, T, X(t))$ equals zero. Thus

$$A'(t, T) + B'(t, T)^T x + B(t, T)^T(\mu_0(t) + \mu_1(t)x)$$

$$+\frac{1}{2}B(t, T)^T(V_0(t) + \sum_{j=1}^{n} V_{1j}(t)x_j)B(t, T) \;\; = \;\; 0$$

where the function $\mu_0(t)$ is $n \times 1$, and $\mu_1(t)$, $V_0(t)$ and the $V_{1j}(t)$ are $n \times n$. Since this is true for all values of $x_1, \ldots, x_n$ this results in a set of $n + 1$ ordinary differential equations. The n equations corresponding to $x_1, \ldots, x_n$ define a set of n simultaneous Riccatti equations. Provided the functions $\mu_0(t)$, $\mu_1(t)$, $V_0(t)$ and the $V_{1j}(t)$ satisfy certain conditions these Riccatti equations can be solved up to any specified $\bar{T} < \infty$. Solution of these equations for a given T and $B(0, T)$ gives us the unique solution for $B(t, T)$. Finally, given T, we can solve for $A(t, T)$

$$A'(t, T) = -B(t, T)^T\mu_0(t) - \frac{1}{2}B(t, T)^T V_0(t)B(t, T).$$

■

References

Black, F., and Karasinski, P. (1991) "Bond and option pricing when short rates are log-normal,"   Financial Analysts Journal July-Aug, 52-29.

Brace, A., Gatarek, D., and Musiela, M. (1997) "The market model of interest-rate dynamics,"   Mathematical Finance 7, 127-155.

Brigo, D., and Mercurio, F. (2001) *Interest Rate Models: Theory and Practice.* Springer: Berlin.

Brody, D.C., and Hughston, L.P. (2001) "Applications of Information Geometry to Interest Rate Theory," in Disordered and Complex Systems, P. Sollich, T. Coolen, L.P. Hughston and R.F. Streater, eds. (AIP), pp. 281-288..

Brody, D.C., and Hughston, L.P. (2002) "Entropy and Information in the Interest Rate Term Structure," Quantitative Finance 2, 70-80.

Cairns, A.J.G. (2004a) "A family of term-structure models for long-term risk management and derivative pricing," Mathematical Finance 14, 415-444.

Cairns, A.J.G. (2004b) *Interest-Rate Models: An Introduction.* Princeton University Press: Princeton. (Forthcoming.)

Chan, K.C., Karolyi, G.A., Longstaff, F.A., and Sanders, A.B. (1992) "An empirical comparison of alternative models of the short-term interest rate.," Journal of Finance 47, 1209-1227.

Cox, J., Ingersoll, J., and Ross, S. (1985) "A theory of the term-structure of interest rates," Econometrica 53, 385-408.

Duffie, D., and Kan, R. (1996) "A yield-factor model of interest rates," Mathematical Finance 6, 379-406.

Flesaker, B., and Hughston, L.P. (1996) "Positive interest," Risk 9(1), 46-49.

Jamshidian, F. (1997) "LIBOR and swap market models and measures," Finance and Stochastics 1, 293-330.

Joshi, M.S., and Rebonato, R. (2003) "A displaced-diffusion stochastic volatility LIBOR market model: motivation, derivation and implementation," Quantitative Finance 3, 458-469.

Rebonato, R. (2002) *Modern Pricing of Interest-Rate Derivatives.* Princeton University Press: Princeton.

Rogers, L.C.G. (1997) "The potential approach to the term-structure of interest rates and foreign exchange rates," Mathematical Finance 7, 157-164.

Rutkowski, M. (1997) "A note on the Flesaker & Hughston model of the term structure of interest rates," Applied Mathematical Finance 4, 151-163.

Vasicek, O. (1977) "An equilibrium characterisation of the term structure," Journal of Financial Economics 5, 177-188.