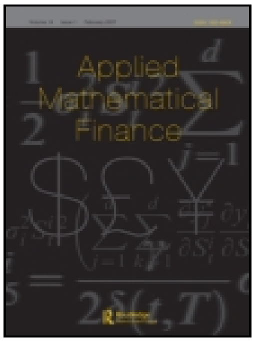

Applied Mathematical Finance

ISSN: 1350-486X (Print) 1466-4313 (Online) Journal homepage: www.tandfonline.com/journals/ramf20

An Affine Two-Factor Heteroskedastic Macro-Finance Term Structure Model

Peter Spreij, Enno Veerman & Peter Vlaar

To cite this article: Peter Spreij, Enno Veerman & Peter Vlaar (2011) An Affine Two-Factor Heteroskedastic Macro-Finance Term Structure Model, Applied Mathematical Finance, 18:4, 331-352, DOI: 10.1080/1350486X.2010.517664

To link to this article: https://doi.org/10.1080/1350486X.2010.517664

Published online: 02 Feb 2011.

Applied Mathematical Finance,
Vol. 18, No. 4, 331–352, September 2011

Routledge
Taylor & Francis Group

An Affine Two-Factor Heteroskedastic Macro-Finance Term Structure Model

PETER SPREIJ*, ENNO VEERMAN* & PETER VLAAR**

**Korteweg-de Vries Institute for Mathematics, Science Park 904, 1098SH Amsterdam, the Netherlands,*
***All Pensions Group, Gustav Mahlerplein 3, 1082MS Amsterdam, the Netherlands*

(Received 20 May 2008; in revised form 27 July 2010)

ABSTRACT *We propose an affine macro-finance term structure model for interest rates that allows for both constant volatilities (homoskedastic model) and state-dependent volatilities (heteroskedastic model). In a homoskedastic model, interest rates are symmetric, which means that either very low interest rates are predicted too often or very high interest rates not often enough. This undesirable symmetry for constant volatility models motivates the use of heteroskedastic models where the volatility depends on the driving factors.*

For a truly heteroskedastic model in continuous time, which involves a multivariate square root process, the so-called Feller conditions are usually imposed to ensure that the roots have non-negative arguments. For a discrete time approximate model, the Feller conditions do not give this guarantee. Moreover, in a macro-finance context, the restrictions imposed might be economically unappealing. It has also been observed that even without the Feller conditions imposed, for a practically relevant term structure model, negative arguments rarely occur.

Using models estimated on German data, we compare the yields implied by (approximate) analytic exponentially affine expressions to those obtained through Monte Carlo simulations of very high numbers of sample paths. It turns out that the differences are rarely statistically significant, whether the Feller conditions are imposed or not. Moreover, economically, the differences are negligible, as they are always below one basis point.

KEY WORDS: Macro-finance models, affine term structure model, expected inflation, *ex ante* real short rate, Monte Carlo simulations

1. Introduction

In this article we propose a two-factor *heteroskedastic* macro-finance affine term structure model (ATSM). A term structure model in general involves one or more driving factors, which are usually assumed unobservable. In the macro-finance models, the driving factors involve macro-economic variables, for instance, the inflation rate, and have a direct economic interpretation. Most macro-finance term structure

---

This article was written while the second and third authors were affiliated with De Nederlandsche Bank. The views in this article are those of the individual authors and do not necessarily reflect official positions of De Nederlandsche Bank or the All Pensions Group.

Correspondence Address: Peter Spreij, Korteweg-de Vries Institute for Mathematics, Science Park 904, 1098SH Amsterdam, the Netherlands. Email: spreij@uva.nl

1350-486X Print/1466-4313 Online/11/040331–22
DOI: 10.1080/1350486X.2010.517664

models encountered in the literature are *homoskedastic*, a constant volatility for the driving factors is assumed, see for instance Campbell and Viceira (2002), Ang and Piazzesi (2003), Ang and Bekaert (2004), Dewachter *et al.* (2004, 2006), Fendel (2005), Bernanke *et al.* (2005), Dewachter and Lyrio (2006), Hördahl *et al.* (2006), Wu (2006) and Rudebusch and Wu (2007).[1] This, however, implies that interest rates are assumed symmetric, which means that either very low interest rates are predicted too often or very high interest rates not often enough. Especially in a low interest rate environment, this characteristic also means that the probability of drawing negative interest rates in a simulation exercise is substantial. Rudebusch (2010) therefore stated: "In the future, developing versions of the affine arbitrage-free model that prevent interest rates from going negative will be a priority". Our article is one of the first attempts for this in the macro-finance context.

The undesirable symmetry for constant volatility models motivated us to include heteroskedasticity by making the volatility dependent on the driving factors. We opt in the first place for a continuous time ATSM driven by a multidimensional Brownian motion. In such a model the volatility involves square roots of affine functions of the state factors. Note that such a model also encompasses the aforementioned homoskedastic models because a constant volatility is a special case of it. These affine models form, as a consequence, a more flexible class than those in which the volatility is constant.

At the same time, the use of affine models naturally imposes restricting conditions on parameter values, because the arguments of the square roots have to be non-negative. These conditions, the *multivariate Feller conditions*, have originally been introduced in Duffie and Kan (1996), see also Piazzesi (2005). Alternatively, they are also known as *admissibility conditions*, see Duffie *et al.* (2003). On a technical note and from a mathematical point of view, these conditions imply more than merely non-negative arguments of square roots. It is known, see Duffie and Kan (1996), that the underlying stochastic differential equations (SDE2) admit strong solutions (this topic is only remotely relevant for this study), and they also imply that closed form expressions for bond prices can be obtained. Violation of these conditions has therefore unappealing consequences, at least from a technical point of view. On the contrary, as we shall show, maintaining these conditions may sometimes result in unattractive implications from an economic macro-finance point of view, as they then describe untenable relations between certain macroeconomic quantities. Therefore, there is a potential conflict between a mathematical requirement and the adopted economic theory.

In practice, one often works with discrete time models, which, for instance, can be obtained by discretizing a continuous time model. If the continuous time model is driven by a Brownian motion, its discrete time counterpart will contain normally distributed error terms. For the affine models that we consider, also in discrete time approximations, the volatility factor will contain square roots of affine functions of the state variables. In a first approach it is therefore natural to impose multivariate Feller conditions again. However, their usefulness is debatable in a discrete time setting. These conditions are insufficient to guarantee that the square roots always have non-negative arguments. Indeed, the standard normally distributed errors, which are used as inputs in these discrete time models, imply that at each time instant there is a positive probability that one or more of the arguments of the square roots become negative,

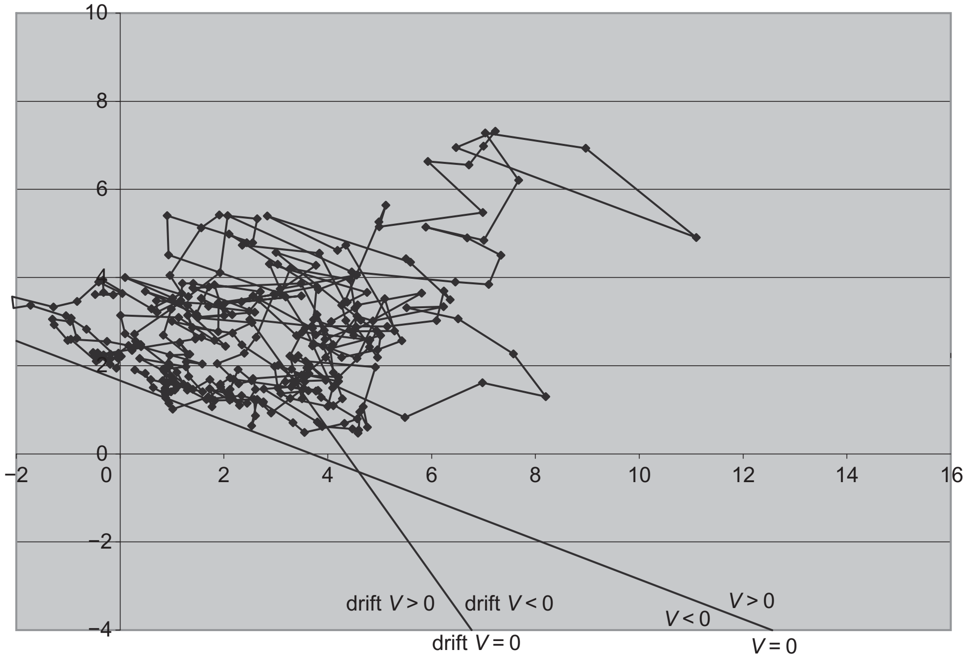

Figure 1. Example of a sample path of a discrete time ATSM for which the Feller conditions are not satisfied.

regardless whether the Feller conditions are satisfied or not.[2,3] Despite of all this, it is still not uncommon to impose the Feller conditions on discrete time models, as they serve as approximations of continuous time models. However, one might also question the necessity of these conditions, as indicated by the following empirical study.

We consider a two-dimensional factor process in discrete time. In Figure 1 we show a typical simulated trajectory, obtained for parameter values that do not satisfy the Feller conditions. In this setting, with a single volatility factor (denoted v), one of the Feller conditions imposes that for every point on the line $v = 0$, the deterministic part of the process (the drift) is such that the volatility becomes positive again. Although it is clear that this condition is necessary in a univariate setting, its significance in a multivariate setting is not obvious as the interaction between the factors limits the part of the line $v = 0$ that is actually approached. We conclude from Figure 1 that although the line $v = 0$ is crossed once, this happens in the area where the drift of the volatility process is positive, an impossibility in a continuous time model. In other words, the fact that v also assumes negative values is because of the discretization of the model, not because the Feller conditions are violated.

Combined with the potential conflict that may arise with certain principles prescribed by the economic theory whether the Feller conditions are imposed, this observation motivates a study that sheds some light on the necessity to impose these conditions.

We will work with a two-dimensional ATSM, with the macro-economic quantities *ex ante* real short rate and expected inflation as state variables.These models allow for state-dependent (heteroskedastic) volatility models as well as a more traditional homoskedastic constant volatility model. Estimation is performed by maximum likelihood for both the homoskedastic model and different types of heteroskedastic affine models, which will be referred to as models with *independent volatilities*, *dependent volatilities* and *proportional volatilities*. Each of these models will be estimated, with and without the Feller conditions imposed. In a pure latent variable model, it is usual to impose these conditions by assuming a *canonical form*, as in Dai and Singleton (2000). In such a canonical form, the volatility factors are equal to some of the state factors. However, we cannot do this for a macro-finance model as ours, because none of the factors can be taken as a volatility factor *a priori*. Therefore, we need and extract explicit parameter restrictions from the Feller conditions for our non-latent variable model. The estimation is complicated by the fact that expected inflation is an unobservable. We use the Kalman filter combined with a likelihood approach to estimate the involved parameters. We compare the obtained results for the different heteroskedastic models to results obtained for a homoskedastic model. We will see that heteroskedastic models outperform a constant volatility model.

Having executed the estimation, we compare two consecutive approaches to validate bond prices. In the first one, we calculate the bond prices directly given the estimated parameters, using the discrete time Riccati equations (thereby ignoring the cut-off of volatility at zero for the affine models). In the second approach, we perform a high number of Monte Carlo simulations of the trajectories of the factors, whereby volatility is restricted to be non-negative. The mean of these simulations gives a second approximation of the bond price. Moreover, we can measure the approximation error with (sampled) confidence intervals. We will show that the differences between Monte Carlo results and the values obtained from the exponential affine formula are almost always negligible, both economically and statistically, whether the Feller conditions are imposed or not. From an economic point of view, the difference in implied yields between the two methods is hardly relevant, as it is at most one basis point. Statistically, the difference is only significantly different from 0 for some maturities for the dependent and independent volatility models without Feller conditions. For the proportional volatility models (and the constant volatility model), there is never a problem.

The rest of this article is organized as follows. In Section 2 we review general model2 and ATSM2 in continuous time, mainly with the aim to set the notation for the sections to follow. Then we discretize a continuous time model and show that the discretized model leads to the same expression for bond prices as the discretized version of formula for bond prices in continuous time. In Section 3 we present and estimate our models. The explicit expressions for the Feller conditions for our model, which is not given in canonical form, are given in Appendix A, which contains further technical derivations. In Section 4 we use the estimated models of Section 3 to price bonds by Monte Carlo methods and compare the obtained results with those obtained by analytic methods. Finally in Section 5 we summarize our findings and draw some conclusions.

2. Affine Term Structure Models

Although we propose an ATSM for *discrete* time, we first discuss *continuous* time models. The reason for this is that the mathematical theory for ATSM 2 was initially developed for continuous time models (Duffie and Kan, 1996), whence in this respect it is natural to regard the equations governing the discrete time model as discretizations of the continuous time equations.

2.1 Short-Rate Term Structure Models

Let us first concisely review some general theory of short-rate term structure models, see Hunt and Kennedy (2000), Musiela and Rutkowski (1997) or Brigo and Mercurio (2006) for details. The formulas below will be used in subsequent sections. We assume that all relevant expressions are well defined.

In a short-rate term structure model, the price $D_{t,T}$ of a zero-coupon bond at time t maturing at T is based on the dynamics of the short-rate r through the formula

$$D_{t,T} = E_{\mathbb{Q}}\left(\exp\left(-\int_t^T r_s ds\right)|\mathcal{F}_t\right), \tag{1}$$

with $\mathbb{Q}$ the risk-neutral measure and $\mathcal{F}_t$ the underlying filtration.

Typically, in a short-rate model one chooses r to be a function of a (possibly multi-dimensional) process X which satisfies an SDE

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t^{\mathbb{Q}},$$

with $W^{\mathbb{Q}}$ a multivariate Brownian motion under the risk-neutral measure $\mathbb{Q}$ and one writes $r_t = r(X_t)$.

Under rather general conditions, there exists a strong solution X to this equation which is Markov. In this case the bond-price can be written as $D_{t,T} = E_{\mathbb{Q}}(\exp(-\int_t^T r(X_s)ds)|X_t) =: F(t, X_t)$ for some function F (with T fixed). If F is smooth enough, then it solves the *fundamental partial differential equation* (PDE), also called *term structure equation* (see Musiela and Rutkowski (1997, Chapter 12) or Vasicek (1977), where the latter terminology was introduced)

$$\frac{\partial}{\partial t}F(t,x) + \mathcal{L}F(t,x) - r(x)F(t,x) = 0, \quad F(T,x) = 1, \tag{2}$$

with

$$\mathcal{L} = \sum_i \mu_i \frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j}(\sigma\sigma^\top)_{ij}\frac{\partial^2}{\partial x_i \partial x_j},$$

the generator of X, where $\sigma^\top$ means the transpose of σ.

The physical measure $\mathbb{P}$ is on $\mathcal{F}_T$ equivalent to the risk-neutral measure $\mathbb{Q}$ and related through a density process L by $L_t = \frac{d\mathbb{P}}{d\mathbb{Q}}|_{\mathcal{F}_t}$. The process L can often be written as an exponential process $\mathcal{E}(Y \cdot W^{\mathbb{Q}})$ for some Y, that is,

$$L_t = \exp\left(\int_0^t Y_s^\top dW_s^{\mathbb{Q}} - \frac{1}{2}\int_0^t Y_s^\top Y_s ds\right), \tag{3}$$

where Y_s is usually called the *market price of risk*. According to Girsanov's theorem $W_t^{\mathbb{P}} = W_t^{\mathbb{Q}} - \int_0^t Y_s ds$ is a $\mathbb{P}$-Brownian motion, see Karatzas and Shreve (1991, Section 3.5) for details on absolutely continuous measure transformations. Using these relations, one can write the SDE for X under the physical measure $\mathbb{P}$:

$$dX_t = (\mu(t, X_t) + \sigma(t, X_t) Y_t)dt + \sigma(t, X_t)dW_t^{\mathbb{P}}. \tag{4}$$

2.2 ATSM2 in Continuous Time

ATSMs are examples of short-rate models and were introduced by Duffie and Kan (1996). In an ATSM the short-rate r is an affine function of X, that is, $r = \delta_0 + \delta^\top X$ for some $\delta_0 \in \mathbb{R}$, $\delta \in \mathbb{R}^n$, and X satisfies under $\mathbb{Q}$ an n-dimensional affine square root SDE

$$dX_t = (aX_t + b)dt + \Sigma\sqrt{v(X_t)}dW_t^{\mathbb{Q}}. \tag{5}$$

Here $W^{\mathbb{Q}}$ is an n-dimensional Brownian motion, $v(X_t)$ is a diagonal matrix with on its diagonal the elements of the vector

$$\mathrm{diag}[v(X_t)] = \alpha + \beta X_t, \tag{6}$$

with $\alpha \in \mathbb{R}^{n\times 1}$, $\beta \in \mathbb{R}^{n\times n}$ (so $v_{ii}(x) = \alpha_i + \beta_i x$, with β_i the i-th row vector of β). We will call these elements *volatility factors* and we write $V_t := v(X_t)$ and $V_{i,t} := v_i(X_t)$. For brevity, we denote by $\sqrt{V_t}$ the matrix with on the diagonal the square roots $\sqrt{V_{i,t} \vee 0}$, that is the square root of the maximum of $V_{i,t}$ and 0. We will also use the notation $V_t \vee 0$ for the diagonal matrix with elements $V_{i,t} \vee 0$. Notice that $(\sqrt{V_t})^2 = V_t \vee 0$. Existence and uniqueness of a strong solution can be established when the volatility factors stay strictly positive. Conditions for the latter are derived in Duffie and Kan (1996) and are often called *multivariate Feller conditions*. We discuss these in Appendix A.1.

The practical benefit of ATSM2 is that bond prices are determined by a closed form expression. Duffie and Kan (1996) showed that for positive volatility factors the term structure Equation (2) is solved by

$$F(t, x) = \exp(A(T - t) + B(T - t)^\top x), \tag{7}$$

for $t \in [0, T]$ and $x \in \mathcal{D} := \{x \in \mathrm{R}^n\colon\ v_i(x) \geq 0, \forall i\}$, where A and B satisfy the Riccati ordinary differential equations (ODEs)[4]

$$A' = b^\top B + \frac{1}{2}\alpha^\top(\Sigma^\top B)^{\odot 2} - \delta_0, \quad A(0) = 0; \tag{8}$$

$$B' = a^\top B + \frac{1}{2}\beta^\top(\Sigma^\top B)^{\odot 2} - \delta, \quad B(0) = 0. \tag{9}$$

Hence the bond prices can be calculated by

$$D_{t,T} = \exp(A(T-t) + B(T-t)^\top X_t),$$

provided the volatility factors stay positive (for which the aforementioned Feller conditions are sufficient).

In ATSM2, it is often desired that the process X also satisfies an affine square root SDE under the physical measure $\mathbb{P}$, which considerably restricts the choice for the market price of risk Y. We only consider the so-called *completely* affine model, which means that we take $Y_t = \sqrt{V_t}\lambda$ with $\lambda \in \mathbb{R}^\kappa$, and we refer to Duffee (2002) (essentially affine models) and Cheridito *et al.* (2007) (extended affine models) for other options. In a completely affine model, the SDE (4) takes the form

$$dX_t = (aX_t + b + \Sigma(\sqrt{V_t})^2\lambda)dt + \Sigma\sqrt{V_t}dW_t^{\mathbb{P}}. \tag{10}$$

Under the condition that $V_t \geq 0$ (elementwise on the diagonal), it holds that $(\sqrt{V_t})^2 = V_t$ and Equation (10) reduces to

$$dX_t = (\widehat{a}X_t + \widehat{b})dt + \Sigma\sqrt{V_t}dW_t^{\mathbb{P}}, \tag{11}$$

with $\widehat{a} = a + \Sigma(\beta \odot \lambda)$, $\widehat{b} = b + \Sigma(\alpha \odot \lambda)$. The affine structure of (5) thus carries over to (11) for positive volatility factors.

2.3 ATSM2 in Discrete Time

In this section we take Equations (1), (5), (7), and (11) as our point of departure and transform them into their discrete time counterparts, using the Euler method (Kloeden and Platen, 1999). Next we investigate whether the resulting equations are consistent with each other.

In order not to complicate notation, we assume a discretization factor equal to 1. We write $P_{n,t}$ for the bond price at time t maturing at time $t+n$ (which corresponds to $D_{t,t+n}$ in continuous time). Basically, all continuous time formulas are translated to discrete time by replacing integrals by sums and substituting Δ for d. By the properties of the Brownian motion, we have $\Delta W_t^{\mathbb{Q}} = W_{t+1}^{\mathbb{Q}} - W_t^{\mathbb{Q}} \sim N(0, I)$ for each t, and all these increments are mutually independent. Therefore, we write $\varepsilon_{t+1}^{\mathbb{Q}}$ instead of $\Delta W_t^{\mathbb{Q}}$, with $\varepsilon_{t+1}^{\mathbb{Q}}$ i.i.d. standard normal variables under the risk-neutral measure $\mathbb{Q}$. For the filtration we choose the natural filtration $\mathcal{F}_t = \sigma(\varepsilon_k^{\mathbb{Q}} : k = 1, \ldots, t)$. We assume that $\mathbb{Q}$ and the physical measure $\mathbb{P}$ on $\mathcal{F}_t$ are related by $d\mathbb{P} = \tilde{\tilde{L}}\mathbb{Q}$, with L the discretized exponential process

$$\tilde{L}_t = \exp\left(\sum_{k=0}^{t-1} \lambda^\top \sqrt{V_k}\varepsilon_{k+1}^{\mathbb{Q}} - \frac{1}{2}\sum_{k=0}^{t-1} \lambda^\top (V_k \vee 0)\lambda\right).$$

In the continuous time case, $W^{\mathbb{P}}$ defined by $dW^{\mathbb{P}} = dW_t^{\mathbb{Q}} - \sqrt{V_t}\lambda dt$ is a Brownian motion under $\mathbb{P}$ according to Girsanov's theorem. Analogously, in discrete time, one can show that the $\varepsilon_t^{\mathbb{P}}$ defined by $\varepsilon_{t+1}^{\mathbb{P}} = \varepsilon_{t+1}^{\mathbb{Q}} - \sqrt{V_t}\lambda$ are i.i.d. standard normal variables under $\mathbb{P}$. Therefore, we replace $dW_t^{\mathbb{P}}$ in Equation (11) with $\varepsilon_{t+1}^{\mathbb{P}}$.

Application of the above substitutions enables us to transform the continuous time model into a discrete time model. The two SDE2 (5) and (11) under $\mathbb{Q}$ and $\mathbb{P}$, respectively, transform into

$$X_{t+1} = (I + a)X_t + b + \Sigma\sqrt{V_t}\varepsilon_{t+1}^{\mathbb{Q}}, \tag{12}$$

$$X_{t+1} = (I + \widehat{a})X_t + \widehat{b} + \Sigma\sqrt{V_t}\varepsilon_{t+1}^{\mathbb{P}}. \tag{13}$$

Note that the existence of a unique strong solution is not an issue here, so in this respect positive volatility factors are not necessary. The bond price formula (1) becomes

$$P_{n,t} = E_{\mathbb{Q}}\Big(\exp(-\sum_{k=0}^{n-1} r_{t+k})|\mathcal{F}_t\Big). \tag{14}$$

Finally, the closed form expression (7) for the bond price in continuous time corresponds to $\tilde{F}(n, X_t) = \exp(A_n + B_n^\top X_t)$ in discrete time, with $n = T - t$ and A_n and B_n the Euler discretizations of the solutions of the ODEs (8) and (9). The latter means that A_n and B_n satisfy the *Riccati recursions*

$$A_{n+1} = A_n + b^\top B_n + \frac{1}{2}\alpha^\top(\Sigma^\top B_n)^{\odot 2} - \delta_0, \qquad A_0 = 0; \tag{15}$$

$$B_{n+1} = (I + a)^\top B_n + \frac{1}{2}\beta^\top(\Sigma^\top B_n)^{\odot 2} - \delta, \qquad B_0 = 0, \tag{16}$$

which are equivalent to

$$A_{n+1} = A_n + (\widehat{b} - \Sigma(\alpha \odot \lambda))^\top B_n + \frac{1}{2}\alpha^\top(\Sigma^\top B_n)^{\odot 2} - \delta_0, \quad A_0 = 0; \tag{17}$$

$$B_{n+1} = (I + \widehat{a} - \Sigma(\beta \odot \lambda))^\top B_n + \frac{1}{2}\beta^\top(\Sigma^\top B_n)^{\odot 2} - \delta, \quad B_0 = 0. \tag{18}$$

Now that we have derived the discrete time equations, it is important to note that it is impossible to prevent the volatility factors $V_{t,i}$ from becoming negative, because the noise variables are normally distributed. So in this respect it is useless to impose the Feller conditions. The possibility of negative volatility factors can lead to consistency problems for our discrete time model. We saw that in continuous time positive volatility factors implied that $(\sqrt{V_t})^2$ was equal to V_t, which enabled us to write Equation (10) as Equation (11), an affine square root SDE under the physical measure. As in discrete time V_t can always become negative, in this case it does *not* hold true that $(\sqrt{V_t})^2 = V_t$, which implies that the dynamics of X given by Equations (12) and (13) are not consistent with each other. Recalling the definition of $\varepsilon^{\mathbb{P}}$, we see that there are two possibilities to solve this problem, either we keep Equation (12) and replace Equation (13) by

$$X_{t+1} = (I + a)X_t + b + \Sigma(V_t \vee 0)\lambda + \Sigma\sqrt{V_t}\varepsilon^{\mathbb{P}}_{t+1}, \tag{19}$$

or we keep Equation (13) and replace Equation (12) by

$$X_{t+1} = (I + \widehat{a})X_t + \widehat{b} - \Sigma(V_t \vee 0)\lambda + \Sigma\sqrt{V_t}\varepsilon^{\mathbb{Q}}_{t+1}. \tag{20}$$

We opt for the latter, because an attractive expression under the physical measure is preferable in view of estimation of the parameters, which is the topic of the next section.

Positive volatility factors in continuous time were also needed to solve the term structure Equation (2) to obtain a closed form expression F for the bond price. There is no discrete time analogue of a term structure equation. However, using induction and the properties of a log-normal variable, we can *algebraically* derive a closed form expression for the bond price in discrete time, and, just as in continuous time, we need positiveness of the volatility factors for this, see Proposition 1, whose proof is deferred to Appendix A.2. It is remarkable is that this leads to the same expression as $\tilde{F}$, the discretization of the closed form expression F in continuous time.

Proposition 1. Let X satisfy Equation (12). Then for $P_{n,t}$ given by Equation (14) it holds that

$$P_{n,t} \geq \tilde{F}(n, X_t) = \exp(A_n + B_n^{\mathrm{T}} X_t), \tag{21}$$

with equality if $V_t \geq 0$ almost surely for all t. The scalars A_n and vectors B_n satisfy the Riccati recursions (15) and (16).

As noted before, the probability that V_t gets negative is positive for all t, so $P_{n,t}$ is *not* equal to $\tilde{F}(n, X_t)$. However, from a heuristic point of view, if V_t gets negative with very small probability, Equations (12) and (20) are “almost” equivalent and the inequality in Equation (21) is “almost” an equality. This suggests that $\tilde{F}$ might be a good approximation for P. Indeed, under the Feller conditions, for the time discretization step converging to zero, the discrete time process X converges to its continuous time counterpart, which produces non-negative volatilities. So in the limit case, Equation (21) becomes an equality. In the remaining sections we implement and estimate the discrete time model for two dimension using real data and we investigate how well in this case $\tilde{F}$ approximates P. This is done by comparing $\tilde{F}$ with Monte Carlo computations for the bond price P, based on Equations (14) and (20).

3. Implementation and Estimation of the Discrete Time ATSM

In this section, we investigate two-factor models with the *ex ante real short-term rate* and *expected inflation* as state variables, the nominal short rate being the sum of these two factors. We use affine models with three different volatility profiles, as specified in Appendix A.1, and a homoskedastic one with constant volatility, as classically used in the macro-finance literature, that serves as a benchmark.

We have chosen to take the *ex ante* real short-term rate and expected inflation as factors, because the short rate is dominated by monetary policy which in turn is affected most by expected inflation. Moreover, the interaction between interest rates and inflation is important for several reasons. For instance, most pension funds have the intention to give indexation, whereas index-linked bonds are hardly available.

3.1 Specification of the Model

Let $X_{1,t}$ denote the *ex ante* real short-term rate at time t and $X_{2,t}$ the expected inflation. We use our dynamic model for quarterly data, so time is measured in quarters. Consequently, r_t as used in the pricing formulas for bonds is given in ordinary fractions per time unit, in our case per quarter. For numerical and readability reasons, however, we want to express X in percentages per year. Therefore, we have $r_t = (X_{1,t} + X_{2,t})/400$.

With respect to inflation, we are primarily interested in the *ex ante* expectation and not so much in past realizations. Let π_{t+1} denote the inflation rate from t to $t+1$ (also in percentages per year), and $X_{2,t}$, its *ex ante* expectation at date t. The observed processes are the short nominal rate r_t and π_t. Because the inflation rate exhibits a seasonal pattern, we also include a seasonal contribution S_t in the model.

Apart from the dynamics of the state process as given in Equation (13), our model is further described by the following equations, which relate the state variables to the observations.

$$r_t = (X_{1,t} + X_{2,t})/400, \tag{22a}$$

$$\pi_{t+1} = X_{2,t} + S_{t+1} + \omega_\pi \sqrt{V_{2,t-1}}\xi_{\pi,t+1}, \tag{22b}$$

$$S_{t+1} = -S_t - S_{t-1} - S_{t-2} + \omega_s \sqrt{V_{2,t-1}}\xi_{s,t+1}, \tag{22c}$$

where $\xi_{\pi,t}$ and $\xi_{s,t}$ are standard normally distributed error terms, which are independent under the physical measure of $\varepsilon_t^{\mathbb{P}}$ (the error term in the equation for X_t).

The data we are using for estimation are the observed longer maturity yields (denoted by $r_{n,t}$, measured in fractions per quarter). These are modelled by the exponential affine expression for the bondprice plus a measurement error, which is assumed to be independently identically distributed among maturities:

$$r_{n,t} := -(A_n + B_n X_t)/n + \left(\nu_0 + \nu_1\sqrt{V_{1,t-1}} + \nu_2\sqrt{V_{2,t-1}}\right)\xi_{n,t}, \tag{23}$$

under the restrictions $\nu_i \geq 0$, and where $\xi_{n,t} \sim N(0, I)$.

Having fully specified the models, we turn to the estimation procedure. A complicating matter is that both factors are not observed. Therefore, the extended Kalman filter (Harvey, 1989) is used to estimate the models.[5] In principle, all parameters can be estimated simultaneously. In practice, however, a one-step procedure tends to lead to unrealistic expected inflation predictions as the best fit for the bond prices is not necessarily achieved for the most realistic expected inflation estimates. As an appropriate modelling of the time series dynamics of interest rates and inflation is considered more important than the lowest measurement error for bond prices, we prefer a two-step procedure. In the first step, the parameters for system (22) are estimated, combined with the dynamics (13) for X_t and its volatility (6). In the second step, the system is augmented by the equations for the long-term yields (23), and we estimate λ and ν, using the Riccati recursions (17) and (18), conditional on the first-step parameters.

3.2 Estimation Results

The models are estimated with quarterly German data over the sample period of 1959 to 2007. To estimate the dynamics between interest rates and inflation correctly, a long sample period is preferred. On the contrary, prices of zero coupon bonds are only available for a relatively short sample period, especially for longer maturities. Therefore, an unbalanced panel was used, with the short rate and inflation data starting in the last quarter of 1959, 1-, 2-, 4-, 7- and 10-year rates starting the third quarter of 1972, the 15-year rate from 1986:02, and the 30-year rate from 1996:01 on.

Table 1 shows the estimation results for the constant volatility and for the state-dependent volatility models without the Feller conditions imposed. As in the latter case only the conditional covariance matrix of the noise terms, which is given by $\Sigma(V_t \vee 0)\Sigma^\top$, is identifiable, we fix $\Sigma_{11} = 1$ in all models, we take $\Sigma_{12} = 0$ in the proportional model and $\Sigma_{22} = 1$ in the independent model. By definition, $\beta_1 = \beta_2$ in the proportional and dependent models. For the constant volatility model, V_t is an identity matrix and we choose a Choleski decomposition of the covariance to identify Σ. As the free estimate of the impact of expected inflation on the real interest rate ($\hat{a}_{12}$) became unrealistically high in the constant volatility model, we restricted this impact to be at most 0.125.

For all four models, the mean real short rate is about 2.4% per year, whereas the mean inflation rate is just over 3%, the values in the first row of Table 1. With respect to the interaction between the short real rate and expected inflation, the lagged response ($\hat{a}$) is in accordance with economic theory. Higher expected inflation leads to higher real rates ($\hat{a}_{12} > 0$), whereas higher real rates depress future inflation ($\hat{a}_{21} < 0$), especially for the constant volatility model. The latter effect is far from significant though. With respect to volatility, both higher real rates and higher expected inflation lead to significantly higher variances, because the elements of β are all positive.

To illustrate the problem with the constant volatility model, Figure 2 compares the empirical cumulative distribution function of the short-term interest rate over our sample with the equilibrium one according to both the constant volatility and the proportional volatility models without Feller conditions. The constant volatility model results in a symmetric distribution for the state variables, and thereby for the short-term interest rate. In reality, however, the distribution of interest rates is asymmetric, whereby extremely low values are less likely than extreme high ones. Because of the

Table 1. Estimation results without the Feller conditions imposed.

	Constant volatilities	Proportional volatilities	Dependent volatilities	Independent volatilities
$\left(-\hat{a}^{-1}\hat{b}\right)^{\top}$	$\begin{bmatrix} 2.39 & 3.06 \\ (5.5) & (5.3) \end{bmatrix}$	$\begin{bmatrix} 2.36 & 3.05 \\ (2.9) & (4.5) \end{bmatrix}$	$\begin{bmatrix} 2.39 & 3.03 \\ (3.0) & (6.3) \end{bmatrix}$	$\begin{bmatrix} 2.33 & 3.12 \\ (2.0) & (5.2) \end{bmatrix}$
$I+\hat{a}$	$\begin{bmatrix} 0.865 & 0.125 \\ (16.3) & (-) \\ -0.073 & 0.972 \\ (1.5) & (23.2) \end{bmatrix}$	$\begin{bmatrix} 0.926 & 0.087 \\ (26.5) & (1.7) \\ -0.002 & 0.938 \\ (0.1) & (24.0) \end{bmatrix}$	$\begin{bmatrix} 0.948 & 0.111 \\ (26.3) & (2.8) \\ -0.026 & 0.950 \\ (0.9) & (25.5) \end{bmatrix}$	$\begin{bmatrix} 0.946 & 0.102 \\ (24.4) & (2.2) \\ -0.006 & 0.940 \\ (0.2) & (24.3) \end{bmatrix}$
$\alpha^{\top}$	$\begin{bmatrix} 1 & 1 \\ (-) & (-) \end{bmatrix}$	$\begin{bmatrix} -0.377 & -0.377 \\ (2.3) & (2.3) \end{bmatrix}$	$\begin{bmatrix} -0.373 & -0.165 \\ (12.2) & (1.4) \end{bmatrix}$	$\begin{bmatrix} -0.377 & -0.081 \\ (3.2) & (1.5) \end{bmatrix}$
β	$\begin{bmatrix} 0 & 0 \\ (-) & (-) \\ 0 & 0 \\ (-) & (-) \end{bmatrix}$	$\begin{bmatrix} 0.105 & 0.230 \\ (2.3) & (2.9) \\ 0.105 & 0.230 \\ (2.3) & (2.9) \end{bmatrix}$	$\begin{bmatrix} 0.108 & 0.194 \\ (4.7) & (10.4) \\ 0.108 & 0.194 \\ (4.7) & (10.4) \end{bmatrix}$	$\begin{bmatrix} 0.117 & 0.193 \\ (2.9) & (3.3) \\ 0.015 & 0.091 \\ (1.4) & (2.6) \end{bmatrix}$
Σ	$\begin{bmatrix} 0.665 & 0 \\ (6.5) & (-) \\ -0.096 & 0.461 \\ (0.6) & (3.6) \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ (-) & (-) \\ -0.257 & 0.639 \\ (1.5) & (5.3) \end{bmatrix}$	$\begin{bmatrix} 1 & -0.260 \\ (-) & (1.7) \\ 0.052 & 0.547 \\ (0.5) & (4.2) \end{bmatrix}$	$\begin{bmatrix} 1 & -0.526 \\ (-) & (3.0) \\ 0.041 & 1 \\ (0.3) & (-) \end{bmatrix}$
$\lambda^{\top}$	$\begin{bmatrix} -0.122 & -0.187 \\ (0.5) & (1.5) \end{bmatrix}$	$\begin{bmatrix} 0.105 & -0.129 \\ (0.7) & (1.3) \end{bmatrix}$	$\begin{bmatrix} 0.108 & -0.136 \\ (0.6) & (1.4) \end{bmatrix}$	$\begin{bmatrix} 0.0524 & -0.209 \\ (0.3) & (1.1) \end{bmatrix}$

Note: Estimation sample 1959:IV–2007:II. Absolute two-step consistent *t*-values in parenthesis.

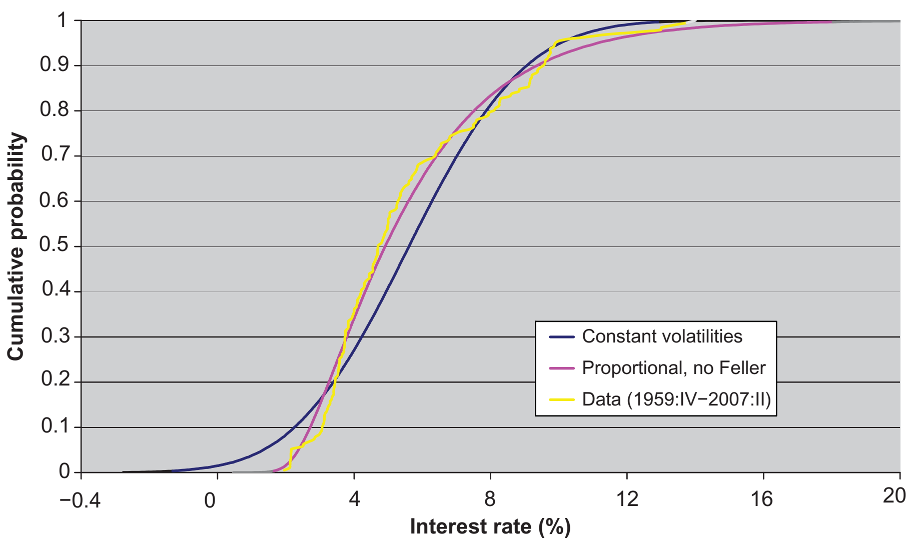

Figure 2. Cumulative distribution function short-term interest rate.

assumed symmetric distribution, a constant volatility model produces a relatively high probability of negative interest rates (here about 1.5%). At the same time, the constant volatility model cannot reproduce the relatively high number of extremely high interest rates. The proportional volatility model (or any of the other heteroskedastic models) is much better capable to reproduce the empirical distribution function and thus shows in this superior behaviour as compared to the homoskedastic constant volatility model.[6]

Table 2 shows the results for the models that are restricted to fulfil the Feller conditions. The resulting constraints on the parameters, when performing the estimation procedure, for each of the three subcases are those as given in Appendix A.1. In the independent volatility model, initially obtained estimates for Σ_{21}, β_{21} and $\hat{a}_{21}$ were practically zero. Therefore, a zero value was subsequently imposed to increase accuracy of the other parameters. In this model, higher inflation now has a negative (though not significant) impact ($\hat{a}_{12} < 0$) on future short-term interest rates, which is contrary to the economic theory, whereas in the previous case when the Feller conditions were not imposed, we found for this coefficient a positive value. In the other models, the impact of inflation on lagged real rates ($\hat{a}_{21} > 0$) is now positive, which is also in contrast with economic theory, see also Appendix A.1 for a theoretical exposition. We conclude that imposition of the Feller conditions may yield a conflict with the realistic economic principles.

Table 2. Estimation results with Feller conditions imposed.

	Proportional volatilities	Dependent volatilities	Independent volatilities
$\left(-\hat{a}^{-1}\hat{b}\right)^{\mathrm{T}}$	$\begin{bmatrix} 2.34 & 3.04 \\ (2.4) & (4.3) \end{bmatrix}$	$\begin{bmatrix} 2.36 & 3.04 \\ (1.9) & (3.3) \end{bmatrix}$	$\begin{bmatrix} 2.91 & 2.83 \\ (2.4) & (3.8) \end{bmatrix}$
$I + \hat{a}$	$\begin{bmatrix} 0.924 & 0.083 \\ (25.0) & (1.5) \\ 0.016 & 0.925 \\ (0.9) & (25.0) \end{bmatrix}$	$\begin{bmatrix} 0.933 & 0.061 \\ (22.2) & (1.2) \\ 0.021 & 0.936 \\ (1.1) & (28.7) \end{bmatrix}$	$\begin{bmatrix} 0.974 & -0.009 \\ (52.4) & (0.4) \\ 0 & 0.958 \\ (-) & (47.9) \end{bmatrix}$
$\alpha^{\top}$	$\begin{bmatrix} -0.412 & -0.412 \\ (2.3) & (2.3) \end{bmatrix}$	$\begin{bmatrix} -0.098 & -0.062 \\ (5.7) & (1.4) \end{bmatrix}$	$\begin{bmatrix} 0.020 & -0.108 \\ (0.7) & (4.5) \end{bmatrix}$
β	$\begin{bmatrix} 0.108 & 0.252 \\ (2.2) & (2.9) \\ 0.108 & 0.252 \\ (2.2) & (2.9) \end{bmatrix}$	$\begin{bmatrix} 0.028 & 0.049 \\ (1.5) & (1.6) \\ 0.028 & 0.049 \\ (1.5) & (1.6) \end{bmatrix}$	$\begin{bmatrix} 0.071 & -0.044 \\ (2.6) & (1.2) \\ 0 & 0.100 \\ (-) & (5.7) \end{bmatrix}$
Σ	$\begin{bmatrix} 1 & 0 \\ (-) & (-) \\ -0.292 & 0.640 \\ (1.8) & (5.6) \end{bmatrix}$	$\begin{bmatrix} 1 & -1.620 \\ (-) & (3.1) \\ 1.116 & 0.915 \\ (1.9) & (2.0) \end{bmatrix}$	$\begin{bmatrix} 1 & 0.615 \\ (-) & (3.4) \\ 0 & 1 \\ (-) & (-) \end{bmatrix}$
$\lambda^{\top}$	$\begin{bmatrix} 0.0050 & -0.124 \\ (0.0) & (1.1) \end{bmatrix}$	$\begin{bmatrix} -0.153 & 0.667 \\ (1.3) & (1.6) \end{bmatrix}$	$\begin{bmatrix} -0.397 & -0.125 \\ (1.2) & (0.1) \end{bmatrix}$

Note: Estimation sample 1959:IV–2007:II. Absolute two-step consistent t-values in parenthesis.

4. Monte Carlo Results

Figure 3 shows the approximation errors made by the analytical expressions, in terms of yields, for each of the seven cases as presented in Tables 1 and 2. The Monte Carlo

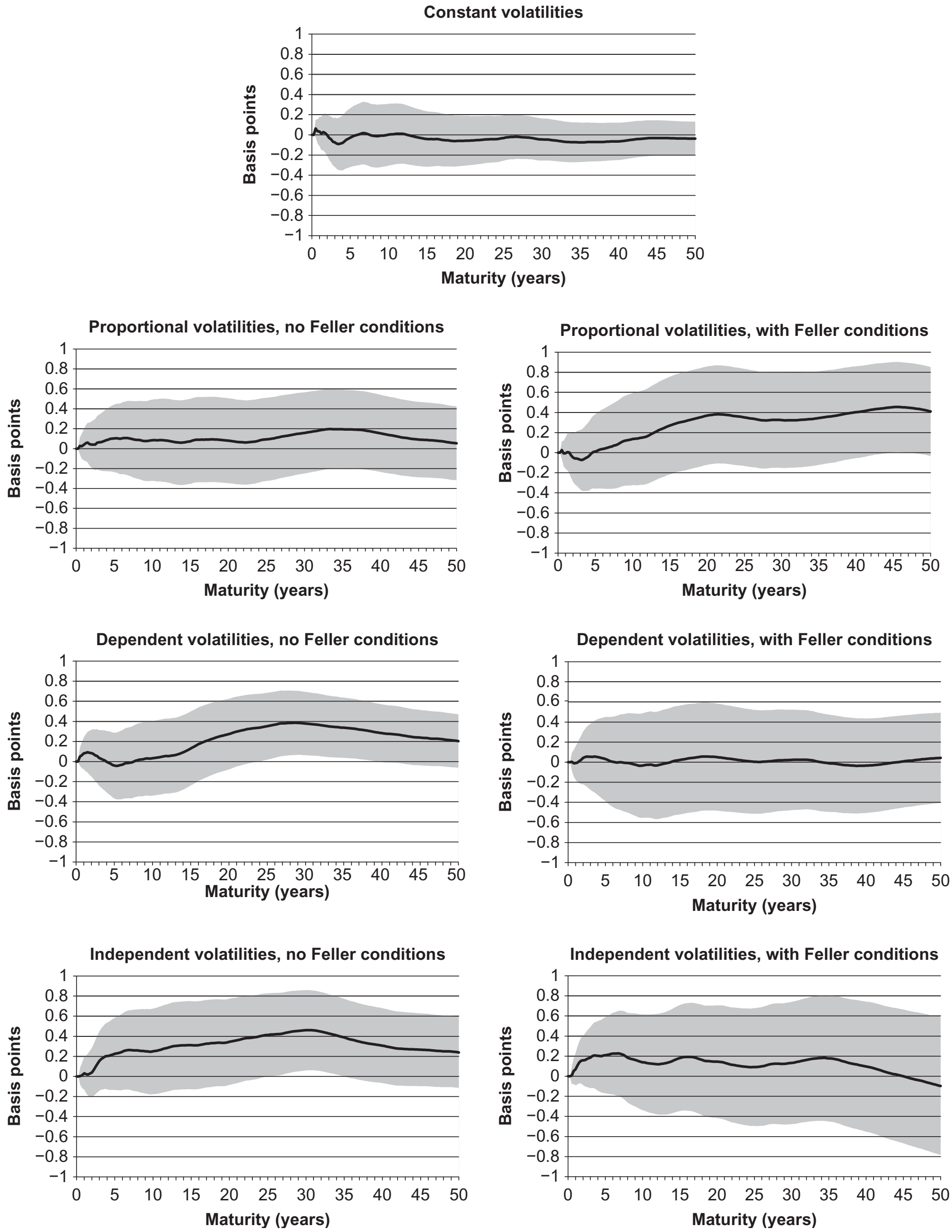

Figure 3. Mean and 99% confidence interval of the difference between simulated and analytical yields if starting state variables are in equilibrium.

simulations are based on one million sample paths (containing 200 quarters) for the state variables. The yields are computed assuming that the initial state variables are at their equilibrium values, which were a real short rate of about 2.4% and expected inflation of just over 3%.

For the constant volatility model, the analytical expression does not contain an approximation, so the result should be very accurate. In this respect, the one million sample paths seem to be sufficient as the 99% confidence interval is less than half a basis point. For the other models, ignoring the fact that volatilities are cut-off at zero does not seem to be very important, though the confidence bands are somewhat larger. The 99% confidence band for the maximum approximation error in terms of yields stays still within plus and minus one basis point (0.01%) for all models. It does not make much difference whether the Feller conditions are imposed (second column) or not (first column). Zero is almost always included in the confidence band, except for some maturities for the dependent and independent volatility models without Feller conditions imposed. For the proportional volatility model, there is never a problem, whether the Feller conditions are imposed or not.

It might be the case that these good approximations are due to the fact that the approximation errors are calculated for the equilibrium yield curve. If the initial state variables imply a volatility closer to zero, ignoring the cut-off at $v = 0$ might be more serious. Therefore, we also calculated the approximation errors for those state variables for which volatility was the lowest in the past. Figure 4 shows the worst result we found. Indeed, for maturities up to 15 years, the simulated yields are significantly higher than the analytical ones. The reason is that for the initial state variables, the volatility is cut-off at zero. As the state variables evolve according to Equation (20),

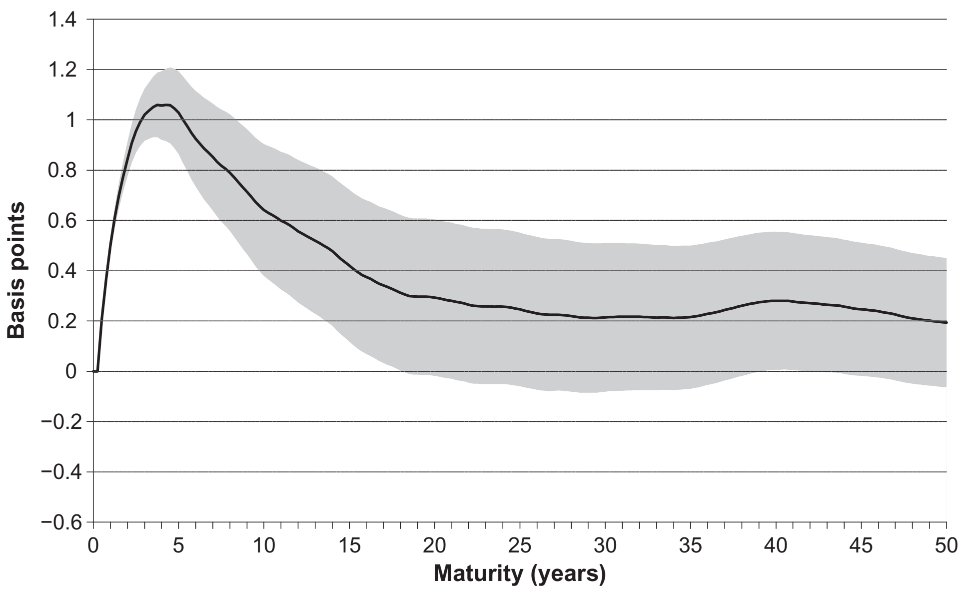

Figure 4. Mean and 99% confidence interval of the difference between simulated and analytical yields. Dependent volatility model without Feller conditions, starting from a volatility of zero.

whereas the formulas underlying the formulas assume Equation (12), systematic differences arise. Moreover, as the simulated yields are almost deterministic for short maturities, the confidence band is extremely small. In economic terms, the approximation error is still negligible though (at most one basis point).

Finally, as the Feller conditions do not guarantee positive volatilities in the discrete time model, imposing them does not preclude statistically significant approximation errors from arising either. Indeed for both the dependent and independent volatility models with Feller conditions, we found starting conditions for which significant negative approximation errors for maturities up to 17 years occur. However, as the maximum magnitude of these errors is at most 0.5 basis point, the economic relevance is again negligible.

5. Conclusions

We have proposed a two-factor ATSM, where the factors (*ex ante* real short-term interest rate and expected inflation) are modelled as a square root process. These models allow for state-dependent volatilities as well as constant volatilities. As we want to allow volatilities to depend on both state variables, the models are not estimated in canonical form.

The Feller conditions, restrictions on the model parameters, are imposed on a continuous time multivariate square root process to ensure that the roots have non-negative arguments. For a discrete time approximate model, the Feller conditions loose part of their relevance. Because the noise involves standard normal errors, there is always a positive probability that arguments of square roots become negative. Nevertheless, keeping in mind the idea that a discrete time model is an approximation of a continuous time model, it is natural to still impose the Feller conditions. On the contrary, it has also been observed that even without the Feller conditions imposed, for a practically relevant model, negative arguments rarely occur. Because the models are not given in canonical form, we also explicitly presented the Feller conditions for square root models not in canonical form.

Three different models with time-varying volatilities have been used that have been referred to as models with proportional, dependent and independent volatilities, either with or without the Feller conditions on the parameters. For comparison reasons, a traditional homoskedastic model is also analysed. The parameters of each of the underlying models have been estimated using quarterly German data. The restrictions involved in imposing the Feller conditions resulted in unappealing economic results. In the proportional and dependent volatility models, the restrictions imply a positive impact of interest rates on inflation, whereas in the independent volatility model, inflation now leads to lower interest rates. Both elements are contrary to the accepted economic theory.

For these seven cases, we have compared the resulting yields that are obtained either by (approximate) analytic exponentially affine expressions or through Monte Carlo simulations of very high numbers of sample paths. It turned out that the approximation errors in analytical yields were rarely statistically significant, and never economically relevant, as they were always below one basis point. In particular, a proportional volatility heteroskedastic model without the Feller conditions imposed already gave very good results, significantly outperforming a traditional constant volatility model.

Notes

[1]An exception is Spencer (2004), who specifies a 10-factor model for the US yield curve, including one heteroskcedasticity factor that is a linear combination of several macroeconomic variables.

[2]This has already been observed in Backus *et al.* (2001, p. 290) (one of the first papers with an affine model in discrete time) for a one-dimensional process, although curiously enough, in the same paper it is claimed that the multivariate Feller conditions are sufficient for non-negative arguments.

[3]An alternative would be to assume a Poisson mixture of Gamma distributions (Dai *et al.* 2005) for the volatility factor instead. For a macro-finance model, this is problematic, however, as volatility is an unknown linear function of the underlying data. Imposing the volatility factor to be equal to one of the driving factors (for instance, expected inflation) has the disadvantage that this factor is not allowed to be influenced by the other factors (for instance, the real short-term interest rate).

[4]The Hadamard product $\odot$ denotes entry-wise multiplication, that is, $(v \odot w)_{ij} = v_{ij} w_{ij}$ for $m \times n$ matrices v and w. We abbreviate $v \odot v$ to $v^{\odot 2}$. Furthermore, we use the Hadamard product also when v is an m-dimensional vector (instead of an $m \times n$ matrix) and w an $m \times n$ matrix (and vice versa), so $(v \odot w)_{ij} = v_i w_{ij}$ (or $(v \odot w)_{ij} = v_{ij} w_i$ in the other case).

[5]The extension is due to the variance equation that includes state variables. Consequently, the true variance process is not known exactly but has to be estimated as well. The resulting inconsistency does not seem to be very important, though in short samples the mean reversion parameters are often biased upwards, see Lund (1997), Duan and Simonato (1999), De Jong (2000), Bolder (2001), Chen and Scott (2003), Duffee and Stanton (2004) and De Rossi (2006).

[6]We used the estimated proportional volatility model to simulate the trajectory of Figure 1.

[7]In Duffie and Kan (1996) it is assumed that $c = 0$, but it is not hard to see that we can take $c > 0$ as well.

References

Ang, A. and Bekaert, K.-G. (2004). The term structure of real rates and expected inflation, CEPR Discussion Paper 4518, Columbia University, New York.

Ang, A. and Piazzesi, M. (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables, *Journal of Monetary Economics*, 50(4), pp. 745–787.

Backus, D., Foresi, S. and Telmer, C. (2001, February). Affine term structure models and the forward premium anomaly, *Journal of Finance*, 56(1), pp. 279–304.

Bernanke, B., Reinhart, V. and Sack, B. (2005). Monetary policy alternatives at the zero bound: an empirical assessment, *Brookings Papers on Economic Activity*, 2, pp. 1–78.

Bolder, D.-J. (2001). Affine term-structure models: theory and implementation, Working Paper 2001–15, Bank of Canada.

Brigo, D. and Mercurio, F. (2006). *Interest Rate Models – Theory and Practice*, *With Smile, Inflation and Credit*, 2nd ed., Springer Finance (Berlin: Springer-Verlag).

Campbell, J. Y. and Viceira, L. M. (2002). Strategic Asset Allocation, Clarendon Lectures in Economics (Oxford: Oxford University Press).

Chen, R.-R. and Scott, L. (2003). Multi-factor Cox-Ingersoll-Ross models of the term structure: estimates and tests from a Kalman filter model, *Journal of Real Estate Finance and Economics*, 27(2), pp. 143–172.

Cheridito, P., Filipovic, D. and Kimmel, R.-L. (2007). Market price of risk specifications and affine models: theory and evidence, *Journal of Financial Economics*, 83(1), pp. 123–170.

Dai, Q., Le, A. and Singleton, K.-J. (2010). Discrete-time dynamic term structure models with generalized market prices of risk, *Review of Financial* Studies, 23, pp. 2184–2227.

Dai, Q. and Singleton, K.-J. (2000). Specification analysis of affine term structure models, *Journal of Finance*, 55(5), pp. 1943–1978.

De Jong, F. (2000). Time-series and cross-section information in affine term structure models, *Journal of Business and Economic Statistics*, 18, pp. 300–314.

De Rossi, G. (2006). Unit Roots and the Estimation of Multifactor Cox-Ingersoll-Ross Models, Manuscript (Cambridge: Cambridge University).

Dewachter, H. and Lyrio, M. (2006). Macro factors and the term structure of interest rates, *Journal of Money, Credit, and Banking*, 38(1), pp. 119–140.

Dewachter, H., Lyrio, M. and Maes, K. (2004). The effect of monetary unification on German bond markets, *European Financial Management*, 10(3), pp. 487–509.

Dewachter, H., Lyrio, M. and Maes, K. (2006). A joint model for the term structure of interest rates and the macroeconomy, *Journal of Applied Econometrics*, 21, pp. 439–462.

Duan, J.-C. and Simonato, J.-G. (1999). Evaluating an alternative risk preference in affine term structure models, *Review of Quantitative Finance and Accounting*, 13, pp. 111–135.

Duffee, G.-R. (2002). Term premia and interest rate forecasts in affine models, *Journal of Finance*, 57(1), pp. 405–443.

Duffee, G.-R. and Stanton, R.-H. (2004). Estimation of Dynamic Term Structure Models, Manuscript (Haas School of Business).

Duffie, D., Filipovic, D. and Schachermayer, W. (2003). Affine processes and applications in finance, *The Annals of Probability*, 13, pp. 984–1053.

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

Fendel, R. (2005). An affine three-factor model of the German term structure of interest rates with macroeconomic content, *Applied Financial Economics Letters*, 1(3), pp. 151–156.

Harvey, A.-C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter* (Cambridge: Cambridge University Press).

Hördahl, P., Tristani, O. and Vestin, D. (2006). A joint econometric model of macroeconomic and term-structure dynamics, *Journal of Econometrics*, 131(1–2), pp. 405–444.

Hunt, P. and Kennedy, J. (2000). *Financial Derivatives in Theory and Practice* (Chichester: Wiley Series in Probability and Statistics).

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

Kloeden, P. E. and Platen, E. (1999). Numerical Solution of Stochastic Differential Equations (Berlin: Springer).

Lund, J. (1997). Econometric analysis of continuous-time arbitrage-free models of the term structure of interest rates, Working Paper, Aarhus School of Business.

Musiela, M. and Rutkowski, M. (1997). Martingale Methods in Financial Modelling, Volume 36 of Applications of Mathematics (New York and Berlin: Springer-Verlag).

Piazzesi, M. (2005). Affine term structure models. In: Y. Aït-Sahalia and L.-P. Hansen (Eds.), *Handbook of Financial Econometrics* (Amsterdam: Elsevier-Science).

Rudebusch, G.-D. (2010). Macro-finance models of interest rates and the economy, Working Paper 2010-01, Federal Reserve Bank of San Francisco.

Rudebusch, G.-D. and Wu, T. (2007). Accounting for a shift in term structure behavior with no-arbitrage and macro-finance models, *Journal of Money, Credit, and Banking*, 39(2–3), pp. 395–422.

Spencer, P. (2004). Affine macroeconomic models of the term structure of interest rates: the US treasury market 1961–99, Discussion Papers in Economics 2004/16, The University of York.

Vasicek, O.-A. (1977). An equilibrium characterization of the term structure, *Journal of Financial Economics*, 5, pp. 177–188.

Wu, T. (2006). Macro factors and the affine term structure of interest rates, *Journal of Money, Credit, and Banking*, 38(7), pp. 1847–1875.

Appendix A: Technical Results

A.1 Multivariate Feller Conditions in Two Dimensions

In this section we present the multivariate Feller conditions as given in Duffie and Kan (1996) and work them out in two dimensions.

Proposition 2. (Duffie and Kan, 1996). *Let X be a solution to the affine square root SDE (5). Then $X_t \in \mathcal{D}, \forall t \geq 0$ holds almost surely under $\mathbb{Q}$ if the multivariate Feller conditions hold, that is for all i, j we have*[7]

$$\beta_i \Sigma^j = 0 \text{ or } v_i = v_j + c \text{ for some } c \geq 0, \tag{A1}$$

$$\beta_i(ax + b) > \frac{1}{2}\beta_i \Sigma\Sigma^\top \beta_i^\top \text{ for all } x \in \partial\mathcal{D}^i. \tag{A2}$$

where we write $\mathcal{D} := \{x \in \mathbb{R}^n : v_i(x) > 0, i = 1, \ldots, n\}$, $\partial\mathcal{D}^i := \{x \in \mathbb{R}^n : v_i(x) = 0, v_j(x) \geq 0, \forall j \neq i\}$ *and* Σ^j *for the j-th column vector of* Σ.

In a pure latent variable model, the Feller conditions can be imposed by assuming a *canonical form* for SDE (5), as shown by Dai and Singleton (2000). In such a canonical form the volatility factors are equal to some of the state factors. However, we cannot do this for macro-finance models as requiring one of the factors to be the volatility factor is overly restrictive. Therefore, we need to extract explicit parameter restrictions from the Feller conditions to impose them for a non-latent variable model. The resulting expressions for two dimension are given below. In the empirical analysis of Section 3, these expressions are crucial.

Let X be a solution to the SDE (5) for $n = 2$. We distinguish three cases: *proportional (linear dependent)* volatilities, *linearly dependent but non-proportional* volatilities and *linearly independent volatilities*. The first case is characterized by $v_2 = kv_1$ for some $k \geq 0$, the second case corresponds to $v_2 = kv_1 + c$ with $k \geq 0$, $c > 0$ and in the third case one has $\det \beta \neq 0$. For the first and second case we take $k = 1$, that is, k is absorbed in Σ, so that we can apply the above proposition. For future reference we also introduce $\gamma_1 = (-\beta_{12}, \beta_{11})$ and $\gamma_2 = (\beta_{22}, -\beta_{21})$, where the β_{ij} are the elements of the matrix β.

Proportional volatilities: As we take $v_2 = kv_1$ with $k = 1$, condition (A1) is automatically satisfied. Hence, for Proposition 2 to hold, we only have to impose (A2). Note that $x \in \partial\mathcal{D}^1 = \partial\mathcal{D}^2$ if and only if $x = -\frac{\alpha_1}{|\beta_1|^2}\beta_1^\top + y\gamma_1^\top$, for some $y \in \mathbb{R}$, where $|\beta_1|$ denotes the Euclidean norm of the vector β_1. Equation (A2) for $i = 1$ becomes $\beta_1(-\frac{\alpha_1}{|\beta_1|^2}a\beta_1^\top + ya\gamma_1^\top + b) > \frac{1}{2}\beta_1\Sigma\Sigma^\top\beta_1^\top$ for all $y \in \mathbb{R}$. This reduces to the following set of conditions

$$-\frac{\alpha_1}{|\beta_1|^2}\beta_1 a\beta_1^\top + \beta_1 b > \frac{1}{2}\beta_1\Sigma\Sigma^\top\beta_1^\top, \tag{A3}$$

$$\beta_1 a\gamma_1^\top = 0. \tag{A4}$$

It is worth noting that the latter condition is in contrast with certain economic principles for the two-factor model of interest rate and inflation (as presented in Section 3). These principles are as follows:

- $a_{12} > 0$: the interest rate will be raised by the Central Bank when the inflation gets too high
- $a_{21} < 0$: prices will fall when the interest rate is high, causing a decrease of the inflation
- $\beta_{11} > 0, \beta_{12} > 0$: the market becomes more volatile when both the values of interest rates and inflation increase.

Under these conditions together with the empirical observation that $0 > a_{22} > a_{11}$ (i.e. real interest rates display stronger mean reversion than expected inflation, see the estimation results in the second column of Table 1), it holds that

$$\beta_1 a \gamma_1^\top = a_{12}\beta_{11}^2 - a_{21}\beta_{12}^2 + (a_{22} - a_{11})\beta_{11}\beta_{12} > 0,$$

which indeed contradicts (A4).

Dependent but unproportional volatilities: In this case $v_2 = v_1 + c$, with $c > 0$. Then condition (1) is automatically satisfied for $i = 2, j = 1$, but for $i = 1, j = 2$ we have to impose the extra condition

$$\beta_1 \Sigma^2 = 0. \tag{A5}$$

Note that $\partial\mathcal{D}^2 = \emptyset$, so for condition (A2) we only have to consider the case $i = 1$. The analysis is completely the same as for the case of proportional volatilities. Hence the conditions of Proposition 2 are equivalent to the set of conditions (A3), (A4) and (A5).

Independent volatilities: Suppose $\det \beta \neq 0$, then β^{-1} exists. Obviously neither $v_2 = v_1 + c$ nor $v_1 = v_2 + c$ holds true for some positive c, so for condition (A1) to hold, we need to impose the restrictions

$$\beta_1 \Sigma^2 = 0, \tag{A6}$$

$$\beta_2 \Sigma^1 = 0. \tag{A7}$$

Note that $x \in \partial\mathcal{D}^1$, respectively $x \in \partial\mathcal{D}^2$, if and only if $\alpha + \beta x \in \{0\} \times \mathbb{R}_{\geq 0]}$, respectively, $\alpha + \beta x \in \mathbb{R}_{\geq 0} \times \{0\}$. Hence condition (A2) is satisfied if and only if

$$\beta_1\left(a\beta^{-1}\left(\begin{pmatrix}0\\w\end{pmatrix} - \alpha\right) + b\right) > \frac{1}{2}\beta_1\Sigma\Sigma^\top\beta_1^\top, \quad \text{for all } w \geq 0, \tag{A8}$$

$$\beta_2\left(a\beta^{-1}\left(\begin{pmatrix}v\\0\end{pmatrix} - \alpha\right) + b\right) > \frac{1}{2}\beta_2\Sigma\Sigma^\top\beta_2^\top, \quad \text{for all } v \geq 0. \tag{A9}$$

As

$$\beta^{-1} = \frac{1}{\det\beta}\begin{pmatrix} \beta_{22} & -\beta_{12} \\ -\beta_{21} & \beta_{11} \end{pmatrix} = \frac{1}{\det\beta}\begin{pmatrix} \gamma_2^\top & \gamma_1^\top \end{pmatrix},$$

we can reduce the restrictions (A8) and (A9) to

$$w\frac{\beta_1 a\gamma_1^\top}{\det\beta} + \beta_1 b - \beta_1 a\beta^{-1}\alpha > \frac{1}{2}\beta_1\Sigma\Sigma^\top\beta_1^\top, \quad \text{for all } w \geq 0.$$

$$v\frac{\beta_2 a\gamma_2^\top}{\det\beta} + \beta_2 b - \beta_2 a\beta^{-1}\alpha > \frac{1}{2}\beta_2\Sigma\Sigma^\top\beta_2^\top, \quad \textit{for all } v \geq 0.$$

These hold true if and only if

$$\frac{\beta_1 a\gamma_1^\top}{\det\beta} \geq 0, \tag{A10}$$

$$\frac{\beta_2 a\gamma_2^\top}{\det\beta} \geq 0, \tag{A11}$$

$$\beta_1 b - \beta_1 a\beta^{-1}\alpha > \frac{1}{2}\beta_1\Sigma\Sigma^\top\beta_1^\top, \tag{A12}$$

$$\beta_2 b - \beta_2 a\beta^{-1}\alpha > \frac{1}{2}\beta_2\Sigma\Sigma^\top\beta_2^\top. \tag{A13}$$

The first two are necessary because v and w can be chosen arbitrarily large, while the latter two follow by choosing $v = w = 0$. In conclusion we can say that the requirements in Proposition 2 are met, if the conditions (A6), (A7) and (A10)–(A13) hold.

It is worth noting that $X_t \in \mathcal{D}, \forall t \in [0, T]$ holds almost surely under $\mathbb{Q}$ if and only if it holds almost surely under $\mathbb{P}$, by the equivalence of $\mathbb{Q}$ and $\mathbb{P}$. Furthermore, X solves (5) if and only if it solves (10), and under the conditions of the proposition it also solves (11). Hence one can rephrase the conditions of the proposition by using the parameters of (11) instead of those of (5), which gives the alternative to (A2), but *under* (A1) equivalent, condition

$$\beta_i(\widehat{a}x + \widehat{b}) > \frac{1}{2}\beta_i\Sigma\Sigma^\top\beta_i^\top, \qquad \text{for all } x \in \partial\mathcal{D}^i. \tag{A14}$$

Consequently, under $\mathbb{P}$ the Feller conditions are also fulfilled under restrictions (A3)–(A13), with a and b replaced by $\hat{a}$ and $\hat{b}$.

A.2 Proof of Proposition 1

Proof We give a proof by induction. For $n = 0$ it holds that $P_{n,t} = P_{0,t} = 1$, so the statement holds true with $A_0 = B_0 = 0$. Now suppose $P_{n-1,t} \geq \exp(A_{n-1} + B_{n-1}^\top X_t)$ for all t and for a certain $n \in \mathbb{N}$. We write

$$P_{n,t} = \mathrm{E}_{\mathbb{Q}}\left[\exp\left(-\sum_{k=0}^{n-1} r_{t+k}\right)\middle|\mathcal{F}_t\right] = \mathrm{E}_{\mathbb{Q}}\left[P_{n-1,t+1}e^{-r_t}|\mathcal{F}_t\right],$$

and use the induction hypothesis to get

$$\begin{aligned}
P_{n,t} &\geq \mathrm{E}_{\mathbb{Q}}[\exp(A_{n-1} + B_{n-1}^\top X_{t+1} - \delta_0 - \delta^\top X_t)|\mathcal{F}_t] \\
&= \mathrm{E}_{\mathbb{Q}}[\exp(A_{n-1} + B_{n-1}^\top((I+a)X_t + b + \Sigma\sqrt{V_t}\varepsilon_{t+1}^{\mathbb{Q}}) - \delta_0 - \delta^\top X_t)|\mathcal{F}_t] \\
&= \exp(A_{n-1} + B_{n-1}^\top b - \delta_0 + ((I+a)^\top B_{n-1} - \delta)^\top X_t)\,\mathrm{E}_{\mathbb{Q}}[\exp(B_{n-1}^\top \Sigma\sqrt{V_t}\varepsilon_{t+1}^{\mathbb{Q}})|\mathcal{F}_t] \\
&\geq \exp(A_{n-1} + B_{n-1}^\top b - \delta_0 + ((I+a)^\top B_{n-1} - \delta)^\top X_t + \frac{1}{2}B_{n-1}^\top \Sigma V_t \Sigma^\top B_{n-1}) \\
&= \exp(\underbrace{A_{n-1} + B_{n-1}^\top b - \delta_0 + \frac{1}{2}\alpha^\top(\Sigma^\top B_{n-1})^{\odot 2}}_{A_n} + (\underbrace{(I+a)^\top B_{n-1} - \delta + \frac{1}{2}\beta^\top(\Sigma^\top B_{n-1})^{\odot 2}}_{B_n})^\top X_t),
\end{aligned}$$

where we have used that (with $\ell = \Sigma^\top B_{n-1}$)

$$\mathrm{E}_{\mathbb{Q}}[\exp(\ell^\top \sqrt{V_t}\varepsilon_{t+1}^{\mathbb{Q}})|\mathcal{F}_t] = \exp(\frac{1}{2}\ell^\top(V_t \vee 0)\ell) \geq \exp(\frac{1}{2}\ell^\top V_t \ell) \tag{A15}$$

and

$$\begin{aligned}
\ell^\top V_t \ell = \ell^\top(\alpha \odot \ell) + \ell^\top(\beta X_t \odot \ell) = \alpha^\top(\ell \odot \ell) + (\beta X_t)^\top(\ell \odot \ell) &= \alpha^\top \ell^{\odot 2} + X_t^\top \beta^\top \ell^{\odot 2} \\
&= \alpha^\top \ell^{\odot 2} + (\beta^\top \ell^{\odot 2})^\top X_t.
\end{aligned}$$

If $V_t \geq 0$ then the inequality in Equation (A15) is an equality. This proves the assertion.