

ASYMMETRIC SHORT-RATE MODEL WITHOUT LOWER BOUND

Frédéric Vrins, Linqi Wang

REPRINT • 2024 / 05

LFIN

Voie du Roman Pays 34, L1.03.01 B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email: lidam-library@uclouvain.be

<https://uclouvain.be/en/research-institutes/lidam/lfin/publications.html>

Asymmetric short-rate model without lower bound

FRÉDÉRIC VRINS   and LINQI WANG  

[†]LFIN/LIDAM, UCLouvain, Voie du Roman Pays 34, Louvain-la-Neuve B-1348, Belgium

[‡]Department of Decision Science, HEC Montréal, Montréal, QC H3T 2A7, Canada

[§]Faculty of Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge CB3 9DD, UK

(Received 20 September 2021; accepted 2 December 2022; published online 21 December 2022)

We propose a new short-rate process which appropriately captures the salient features of the negative interest rate environment. The model combines the advantages of the Vasicek and Cox–Ingersoll–Ross (CIR) dynamics: it is flexible, tractable and displays positive skewness without imposing a strict lower bound. In addition, a novel calibration procedure is introduced which focuses on minimizing the Jensen–Shannon (JS) divergence between the model- and market-implied forward rate densities rather than focusing on the minimization of price or volatility discrepancies. A thorough empirical analysis based on cap market quotes shows that our model displays superior performance compared to the Vasicek and CIR models regardless of the calibration method. Our proposed calibration procedure based on the JS divergence better captures the entire forward rate distribution compared to competing approaches while maintaining a good fit in terms of pricing and implied volatility errors.

Keywords: Affine short-rate model; Negative interest rates; Jensen–Shannon divergence; Implied density calibration

JEL classifications: C52, C61, E43, G12, G13

1. Introduction

The Global Financial Crisis of 2007–2008 prompted major central banks to cut their policy rates to mitigate the contractionary effects of the crisis on the real economy. Faced with the inability to further decrease their policy rates to support the economic recovery due to the (Zero) Lower Bound constraint, central banks resorted to unconventional monetary policies (UMPs), such as quantitative easing and negative interest rate policies, in an attempt to further stimulate the economic recovery in a low interest rate environment. Research has mainly focused on assessing the effectiveness of UMPs in stimulating the real economy; see e.g. Acharya *et al.* (2019) for an analysis of the real effects of the ECB’s Outright Monetary Transactions (OMT) programme and Heider *et al.* (2019) for a study of the transmission of negative interest rates to the real economy through the bank lending channel.

Fewer papers focus on the implications of the low interest rate environment, which results partly from UMPs implemented by central banks, for the modelling and pricing of interest rate derivatives. Filipović *et al.* (2017) introduce the linear-rational term structure framework for the pricing

of swaps and European swaptions and Filipović and Kitapbayev (2018) extend it to American swaption pricing. The authors’ contributions are, however, limited to an environment with a strict Zero Lower Bound and thus provide limited insights on challenges associated with derivative pricing in a negative interest rate environment. Recchioni *et al.* (2017) show empirically that allowing for negative interest rates within the multi-factor Heston stochastic volatility model of Recchioni and Sun (2016) leads to more accurate pricing both in-sample and out-of-sample for European call and put equity options.

We make a contribution to this research agenda by developing a short-rate model which appropriately captures the salient features of the negative interest rate environment while delivering tractable semi-analytical expressions for the pricing of caplets. Mean-reverting processes such as the Vasicek model (Vasicek 1977) and the Cox–Ingersoll–Ross model (Cox *et al.* 1985; CIR hereafter) are considered as standard benchmarks for the modelling of short rate dynamics. The latter was often favoured before the negative interest rate period as it precludes negative values, in contrast to the Vasicek model which postulates a Gaussian distribution. Nevertheless, short-rate modelling frameworks need to be reviewed to better reflect the recent interest rate environment. The most direct solution would be returning to models exhibiting Gaussian

*Corresponding author. Email: lw711@cam.ac.uk

dynamics, such as the Vasicek model, which do not preclude negative values. However, this approach might be at odds with empirical evidence as interest rates display conditional asymmetry in their dynamics; see, e.g. Bauer and Chernov (2021) who document pronounced variations in US yield conditional skewness over the business cycle using option-implied skewness for future Treasury yields. Square-root dynamics, as found in the CIR model, can provide the desired positive skewness but at the cost of non-negativity, which is not appropriate in a negative rate environment. A straightforward solution is to introduce a constant negative shift to the CIR process. Nevertheless, the resulting model would exhibit a strict lower bound (the shift itself). In addition, although one could treat the shift as an additional parameter to be estimated from the data, this approach does not prescribe how to reliably estimate this lower bound parameter and guidelines from the empirical or theoretical literature are currently scarce.

In light of these limitations, this paper makes four contributions: first, we propose a two-factor model (VaCIR hereafter) which combines the desirable features of the Vasicek and CIR models. We also consider the *deterministic shift* extension of these models proposed by Brigo and Mercurio (2001) which guarantees a perfect fit to the term structure of interest rates. The shifted versions of the Vasicek and CIR models are known respectively as the Hull–White model (Hull and White 1990) and the CIR++ model (Brigo and Mercurio 2001). These two models are then combined to obtain the shifted version of our VaCIR model which we denote as VaCIR++. The term structure literature has previously resorted to similar two-factor models to obtain decompositions of risky interest rates into their risk-free components, modelled by a Vasicek process, and compensations for their exposures to credit and/or liquidity risks, modelled by CIR processes; see, for example Longstaff *et al.* (2005) and Filipović and Trolle (2013). Instead, our purpose is to propose a two-factor approach that would better characterize interest rate dynamics in a negative interest rate environment. The tractability of our framework allows us to obtain semi-analytical expressions for the forward rate density and caplet pricing.

Second, we propose a novel approach to calibrate short-rate models by forward rate density matching. This is achieved by minimizing the distance between the market- and model-implied densities, measured by the Jensen–Shannon (JS) divergence. More precisely, we choose the model parameters such that the associated forward rate distribution, computed under the corresponding forward measure, is the *closest* to the one implied by derivative market prices. The key advantage over the price or volatility mean squared error (MSE hereafter) minimization is that our approach analyses the fit based on the full distribution rather than focusing on a few data points. An accurate calibration of the density tends to ensure a good fit in terms of prices or volatilities. However, the converse is not necessarily true. Our proposed calibration methodology generalizes moment-based approaches which focus on a few key moments of the distribution—see, for example Guillaume and Schoutens (2013). By focusing on the whole density, our approach accounts for *all the moments* and circumvents the challenges associated with selecting

which moments to match and how to aggregate the matching errors of the various moments into a single one. Furthermore, by switching the focus of the calibration approach to the minimization of the relative entropy (here, the JS divergence) between the market- and model-implied densities, our contribution also builds directly on the literature applying information theory to quantitative finance. Brody and Hughston (2002) propose a calibration methodology based on Shannon entropy maximization which they apply to interest rate term structure models. Cont and Tankov (2004) suggest a regularized calibration approach for exponential Lévy processes to stock option prices based on relative entropy. By contrast, our approach can be applied to any tractable short-rate models, such as the Hull–White and CIR++ models, and extensions such as our proposed VaCIR++ model. Moreover, instead of using the discrepancy between the densities as a penalty to amend the traditional pricing error minimization approach as in Cont and Tankov (2004), it serves as the main objective function in our calibration method. Finally, Tavin (2012) introduces a calibration approach for Normal Inverse Gaussian models to asset log-returns based on relative entropy matching of the risk-neutral density implied by the volatility smile. However, our contribution departs from this one by adopting a model-free approach to infer the implied forward rate density from market data; using the arbitrage-free smoothing technique of the implied-volatility surface of Fengler (2009) to ensure that the forward rate density obtained is well specified; and opting for the Jensen–Shannon divergence which is better suited to compare densities with potentially different supports.

Third, we study the evolution of the market-implied EURI-BOR forward rate densities during the recent period of negative interest rates. To this end, we collect a time series of normal volatility curves of caps and apply a stripping procedure to obtain caplet prices. The model-free approach of Breeden and Litzenberger (1978) is then used in combination with the arbitrage-free smoothing technique introduced in Fengler (2009) to get smooth non-negative market implied forward rate densities. These densities are the key input to derivative pricing and will serve as the calibration target in the density matching approach we developed. Two papers propose an analysis related to our empirical study of market-implied forward rate densities: Li and Zhao (2009) study the probability density functions of future US LIBOR rates implied from caps data. Similarly, Trolle and Schwartz (2014) provide stylized facts about higher-order swap rate moments based on the implied density obtained from the swaption cube data for Europe and the USA. However, both papers focus on the period preceding the low interest rate regime and their analysis does not provide relevant insights about the negative interest rate environment.

Fourth, we provide a thorough empirical study investigating the performance of our bivariate VaCIR++ model during the recent period of negative interest rates in the euro area. This is achieved by comparing our model to its single-factor components, the Hull–White and CIR++ models, as well as their two-factor extensions, the G2++ and the CIR2++, respectively. We evaluate the performance of these models when calibrated using both the volatility MSE and density matching criteria. We find that the VaCIR++ model, which

combines the features of the Hull–White and CIR++ models, consistently exhibits the best performance regardless of the target criterion. Our results further demonstrate that the JS divergence calibration approach provides a more robust way of calibrating interest rate models to fixed-income derivative data since it has superior performance in capturing the entire forward distribution, high pricing accuracy and comparable fitting quality in terms of matching implied-volatility curves. Therefore, this approach can then be used to price new, complex and illiquid derivatives with the same underlying forward rate.

The remainder of this paper is organized as follows: Section 2 introduces the modelling framework which combines the Hull–White and CIR++ models into our proposed two-factor VaCIR++ model. Section 3 details the different steps of our calibration algorithm. Section 4 presents the data for our empirical application and the results. Section 5 concludes.

2. Model description

In this section, we focus on analytically tractable homogeneous short-rate models existing in the literature and explain how to extend them to obtain a more flexible and general framework that will better account for the negative interest rate environment. For all three considered cases, we use the *deterministic shift* extension illustrated by Brigo and Mercurio (2001). The idea is that by shifting a latent model with time-homogeneous dynamics with a deterministic function $\phi: \mathbb{R}^+ \rightarrow \mathbb{R}$, one obtains a short-rate model that is analytically tractable and able to perfectly fit the term structure of interest rates. Moreover, the shift function ϕ is directly given by the difference between the instantaneous forward rate curve implied from the market and that generated by the latent process, possibly available in closed form.

We start with two classic short-rate models which will serve as benchmarks. The obvious choice is the Hull–White model. We consider an Ornstein–Uhlenbeck (Gaussian) latent process and shift it with the function ϕ . The other natural candidate is the CIR model. Applying a deterministic shift function to the latter yields the CIR++ model. This model features skewness and is compatible with the negative rates environment thanks to the shift. Note that this model exhibits a strict lower bound up to any time horizon T , given by the minimum of the function ϕ on the interval $[0, T]$. The alternative we propose is a combination of the aforementioned models, in an attempt to address their documented shortcomings while preserving the desired features.

2.1. Baseline framework

We postulate a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$ where $\mathbb{Q}$ is a probability measure and $\mathbb{F} := (\mathcal{F}_t)_{t \geq 0}$ the filtration, satisfying the usual conditions. All processes considered in this paper are assumed to be $\mathbb{F}$ -adapted.

We consider two standard interest-rate models that we take as benchmarks, namely the Vasicek model (noted x hereafter) and the CIR model (noted y hereafter). The short-rate

dynamics are respectively given by

$$dx_t = \kappa_1(\theta_1 - x_t)dt + \sigma_1 dW_t^x, \quad x_0 \in \mathbb{R}, \quad (1)$$

$$dy_t = \kappa_2(\theta_2 - y_t)dt + \sigma_2 \sqrt{y_t} dW_t^y, \quad y_0 \in \mathbb{R}_0^+. \quad (2)$$

where W^x and W^y are two independent Brownian motions under $\mathbb{Q}$.

These models belong to the class of homogeneous one-factor affine diffusions, with dynamics[†]

$$dz_t = (a + bz_t)dt + \sqrt{c + dz_t} dW_t, \quad z_0 > -c/d, \quad \forall d > 0, \quad (3)$$

where W is a Brownian motion. We denote the parameters as $\Psi = (a, b, c, d, z_0)$. Duffie and Kan (1996) provide conditions to ensure strictly positive volatility of multivariate affine diffusions. In the univariate case, these conditions reduce to

$$\forall z_t \text{ s.t. } c + dz_t = 0, \quad d(bz_t + a) > \frac{d^2}{2} \Leftrightarrow ad - bc > \frac{d^2}{2}.$$

Intuitively, this condition ensures that the boundary (i.e. the value of the process associated with the vanishing square-root condition, $z_t = -c/d$) will never be hit if the drift of the process, i.e. $a + bz_t$, is sufficiently positive in the neighbourhood of the boundary. The stochastic differential equations (SDEs) associated with the Vasicek and CIR models are recovered from equation (3) using respectively $\Psi^x := (\kappa_1\theta_1, -\kappa_1, \sigma_1^2, 0, x_0)$ and $\Psi^y := (\kappa_2\theta_2, -\kappa_2, 0, \sigma_2^2, y_0)$.[‡] For the Vasicek model, the condition above is satisfied as $ad - bc = \kappa_1\sigma_1^2 > 0 = \frac{d^2}{2}$. For the CIR model, we have $ad - bc = \kappa_2\theta_2\sigma_2^2 > \frac{\sigma_2^4}{2} = \frac{d^2}{2}$. This is equivalent to the Feller condition $\kappa_2\theta_2 > \frac{\sigma_2^2}{2}$, of which the condition in Duffie and Kan (1996) is a multidimensional extension.

Homogeneous affine diffusions, as in equation (3), are popular due to their analytical tractability which simplifies model calibration. For instance, denoting by $\mathbb{E}$ the expectation under $\mathbb{Q}$, it is well known that for such processes,

$$P_t^z(T) := \mathbb{E} \left[e^{-\int_t^T z_s ds} \middle| \mathcal{F}_t \right] = A^z(T-t) e^{-B^z(T-t)z_t}, \quad (4)$$

where A^z, B^z are deterministic functions satisfying Riccati equations, for which analytical solutions can be found.[§]

When $\mathbb{Q}$ stands for the risk-neutral measure, $\mathbb{F}$ is the information flow available to investors and z depicts the short-rate process, $P_t^z(T)$ agrees with $P_t(T)$, the no-arbitrage price at time t of a risk-free zero-coupon bond with maturity $T \geq t$:

$$P_t(T) := \mathbb{E} \left[\frac{M_t^z}{M_T^z} \middle| \mathcal{F}_t \right] = \mathbb{E} \left[e^{-\int_t^T z_s ds} \middle| \mathcal{F}_t \right] = P_t^z(T), \quad (5)$$

where $M^z := (M_t^z)_{t \geq 0}$ and $M_t^z := e^{\int_0^t z_s ds}$ is the money market account numéraire.

[†] While we focus the exposition on the univariate case, multi-dimensional jump-diffusion models can be considered.

[‡] In these expressions, the constraints on the parameters are as follows: $\kappa_1, \sigma_1, y_0, \kappa_2, \theta_2, \sigma_2$ are positive constants and $2\kappa_2\theta_2 > \sigma_2^2$ (Feller condition).

[§] They are recalled in appendix 1 for $z \in \{x, y\}$.

The existence of analytical expressions for zero-coupon bond prices explains why the Vasicek and CIR models are so popular for interest rate modelling. Furthermore, their deterministic shift extensions make it possible to replicate the observed term structure of discount factors at time t . More specifically, the short rate is expressed as $z_t^\phi = z_t + \phi^z(t)$, $z \in \{x, y\}$. The corresponding zero-coupon bond price takes the form

$$P_t^\phi(T) = \exp \left\{ - \int_t^T \phi^z(s) ds \right\} P_t^z(T). \quad (6)$$

One can then choose the function ϕ^z such that the discount curve implied by the z^ϕ model perfectly agrees with the market discount curve, i.e. $P_t^{z^\phi}(T) = P_t^M(T)$. It can be shown that this choice corresponds to

$$\phi^z(t) = f_0^M(t) - f_0^z(t), \quad (7)$$

$$f_0^x(t) = \left(\frac{\sigma_1^2}{2\kappa_1^2} e^{-\kappa_1 t} + \frac{\kappa_1^2 \theta_1 - \sigma_1^2/2}{\kappa_1^2} \right) (1 - e^{-\kappa_1 t}) + x_0 e^{-\kappa_1 t}, \quad (8)$$

$$f_0^y(t) = \frac{2\kappa_2\theta_2(e^{th} - 1)}{2h + (\kappa_2 + h)(e^{th} - 1)} + y_0 \frac{4h^2 e^{th}}{[2h + (\kappa_2 + h)(e^{th} - 1)]^2}, \quad (9)$$

where $f_t^M(\cdot)$ is the time- t instantaneous forward curve associated with the current market term structure, $f_t^z(\cdot)$ is the corresponding curve implied by the model $z \in \{x, y\}$ and $h := \sqrt{\kappa_2^2 + 2\sigma_2^2}$.

2.2. Proposed extension

Recall that our goal is to design a model that would comply with negative interest rates. As discussed in the introduction, the Vasicek model (x) has the ability to deal with negative values, but at the expense of assuming a Normal distribution for the short rate. The shifted CIR model may comply with the negative interest rate environment whenever the shift is negative. However, as explained above, such a process would exhibit a time-dependent lower bound (the shift function itself) which would be associated with additional challenges on how to reliably estimate a meaningful lower bound on interest rates from available data.

The model we propose aims at capturing skewness without imposing a strict lower bound by combining two latent processes: a Vasicek process and a CIR process. We denote the resulting model VaCIR hereafter. Mathematically, we consider the following dynamics for the short rate in this model:

$$r_t := x_t + y_t, \quad (10)$$

where the Brownian motions W^x, W^y in equations (1) and (2) are assumed to be independent. This assumption preserves the analytical tractability of zero-coupon bond prices

$$P_t^r(T) = P_t^x(T)P_t^y(T)$$

$$= A^x(T-t)A^y(T-t)e^{-[B^x(T-t)x_t + B^y(T-t)y_t]}. \quad (11)$$

Similarly, we extend the combined VaCIR model to perfectly replicate the current observed discount curve (denoted VaCIR++ hereafter), i.e. $r_t^\phi = r_t + \phi^r(t)$, where $\phi^r(t) = \phi^x(t) + \phi^y(t) - f_0^M(t)$. Additionally, the zero-coupon bond is defined in the same way as in equation (6). Introducing correlation between the two latent processes would come at the cost of analytical tractability of A, B since the resulting model would not be part of the affine model class, see appendix 3 for details. One consequence of assuming independence between x and y is that our proposed model might yield unsatisfactory calibration results for the term structure of caplet volatilities exhibiting a large hump shape. As the main focus of our analysis is not the calibration of the market-implied volatility surface, we leave these considerations for future research and refer the interested reader to chapter 4 of Brigo and Mercurio (2007) for more details.

2.3. Forward rate dynamics and densities under the forward measure

Due to the central role played by the forward rate in the price of derivative products (e.g. caps and floors), we derive the dynamics and densities related to the forward rate in the various models[†]

$$F_t^z(T, S) = \frac{1}{\Delta} \left(\frac{P_t^z(T)}{P_t^z(S)} - 1 \right), \quad z \in \{x, y, r\}. \quad (12)$$

For convenience, we formulate the dynamics of $F(T, S) = (F_t(T, S))_{0 \leq t \leq T < S}$ under the S -forward measure, $\mathbb{Q}^S$.

PROPOSITION 1 *Let $t \leq T < S$ and $\Delta := S - T$. Then,*

- (a) *The forward rate $F^r(T, S)$ associated with the model specified in equation (10) reads as*

$$F_t^r(T, S) = \frac{1}{\Delta} \left((1 + \Delta F_t^x(T, S))(1 + \Delta F_t^y(T, S)) - 1 \right). \quad (13)$$

Moreover, let $\sigma_t^x := \sigma_1$, $\sigma_t^y := \sigma_2\sqrt{y_t}$ be the diffusion coefficients associated with models x and y , respectively. For $z \in \{x, y\}$, define $W_t^{z,S} := W_t^z + \int_0^t \zeta_s^{z,S} ds$ with $\zeta_t^{z,S} := \sigma_t^z B^z(S-t)$. Define $\mathbb{Q}^S$ using the random variable

$$\frac{d\mathbb{Q}^S}{d\mathbb{Q}} \Big|_{\mathcal{F}_S} = \frac{P_S^r(S) M_0^r}{P_0^r(S) M_S^r} = \frac{P_S^r(S)}{P_0^r(S) M_S^r}.$$

Then,

- (b) *$F^x(T, S), F^y(T, S)$ and $F^r(T, S)$ are $\mathbb{Q}^S$ -martingales;*
 (c) *The dynamics of the forward rates associated with x and y are given by*

$$dF_t^x(T, S) = \sigma^x(t, F_t^x(T, S)) dW_t^{x,S}, \quad dF_t^y(T, S)$$

[†] Of course, with the deterministic shift extension, the expression for forward rates should also be adjusted for the shift term, which can be rewritten as $F_t^{z^\phi}(T, S) = \frac{1}{\Delta} (e^{-\int_t^S \phi^z(s) ds} \frac{P_t^z(T)}{P_t^z(S)} - 1)$, $z \in \{x, y, r\}$.

$$= \sigma^y(t, F_t^y(T, S)) dW_t^{y,S}, \quad (14)$$

where $W^{x,S}, W^{y,S}$ are independent $\mathbb{Q}^S$ -Brownian motions and

$$\begin{aligned} \sigma^x(t, u) &:= \sigma_1 \left[u + \frac{1}{\Delta} \right] (B^x(S-t) - B^x(T-t)), \\ \sigma^y(t, u) &:= \sigma_2 \left[u + \frac{1}{\Delta} \right] \\ &\quad \times \sqrt{(B^y(S-t) - B^y(T-t)) (G_t(T, S, u))}, \\ G_t(T, S, u) &:= \ln \left(\left[1 + \Delta u \right] \frac{A^y(S-t)}{A^y(T-t)} \right). \end{aligned}$$

Proof See appendix 4. ■

Because these processes are independent, the conditional density of r_t given (x_s, y_s) is obtained by the convolution of the corresponding conditional densities for the Vasicek and CIR processes sampled at t . The $\mathbb{Q}^S$ -conditional densities of x and y can be found in Brigo and Mercurio (2007).

PROPOSITION 2 Let $t \leq T \leq S$. Under $\mathbb{Q}^S$, the density of the spot rate z_T^ϕ conditional on z_t^ϕ , for $z \in \{x, y, r\}$, is given by,

$$f_{z_T^\phi | z_t^\phi}^S(u) = f_{z_T | z_t}^S(u - \phi^z(t)). \quad (15)$$

Note that when $z = r$, the condition $z_T | z_t$ becomes $r_T | (x_t, y_t)$.

(a) The density of x_T conditional on x_t is given by

$$\begin{aligned} f_{x_T | x_t}^S(u) &= \frac{1}{\sigma^x(T-t)} \varphi \left(\frac{u - \mu^x(t, T)}{\sigma^x(T-t)} \right), \\ \mu^x(t, T) &= x_t e^{-\kappa_1(T-t)} + M^S(t, T), \\ \sigma^x(\tau) &= \sqrt{\frac{\sigma_1^2}{2\kappa_1} [1 - e^{-2\kappa_1\tau}]}, \\ M^S(t, T) &= \left(\theta_1 - \frac{\sigma_1^2}{\kappa_1^2} \right) (1 - e^{-\kappa_1(T-t)}) \\ &\quad + \frac{e^{-\kappa_1(S-T)}}{\kappa_1} (\sigma^x(T-t))^2, \end{aligned} \quad (16)$$

where φ is the standard Normal density.

(b) The density of y_T conditional on y_t is given by

$$\begin{aligned} f_{y_T | y_t}^S(u) &= f_{\chi^2(v, \delta(t, T, S)) / q(t, T, S)}(u) \\ &= q(t, T, S) f_{\chi^2(v, \delta(t, T, S))}(q(t, T, S)u), \quad u \geq 0, \end{aligned} \quad (17)$$

where $f_{\chi^2(v, \delta)}$ is the density of a non-central chi-squared random variable with v degrees of freedom and non-centrality parameter δ . We introduced the following notation:

$$q(t, T, S) := 2 \left(\rho(T-t) + \frac{\kappa_2 + h}{\sigma_2^2} + B^y(\Delta) \right),$$

$$\begin{aligned} \delta(t, T, S) &:= \frac{4\rho(T-t)^2 y_t e^{h(T-t)}}{q(t, T, S)}, \\ \rho(\tau) &:= \frac{2h}{\sigma_2^2 (\exp[h\tau] - 1)}, \end{aligned}$$

with $v := 4\kappa_2\theta_2/\sigma_2^2$ and $h := \sqrt{\kappa_2^2 + 2\sigma_2^2}$.

(c) The density of r_T conditional on (x_t, y_t) is given by

$$f_{r_T | (x_t, y_t)}^S(u) = \int_0^\infty f_{x_T | x_t}^S(u-v) f_{y_T | y_t}^S(v) dv. \quad (18)$$

From the expressions in Proposition 2, we can now provide a characterization of the conditional density of the zero-coupon bond price under the S -forward measure.

PROPOSITION 3 Let $t \leq T \leq S$. Under $\mathbb{Q}^S$, the density of the zero-coupon bond price, $P_T^z(S)$, conditional upon z_t^ϕ , for models of $z \in \{x, y, r\}$ with shift extensions, is given by

$$f_{P_T^z(S) | z_t^\phi}^S(u) = e^{\int_t^T \phi^z(s) ds} f_{P_T^z(S) | z_t}^S \left(u e^{\int_t^T \phi^z(s) ds} \right), \quad u \geq 0. \quad (19)$$

Note that when $z = r$, the condition $P_T^z(S) | z_t$ becomes $P_T^r(S) | (x_t, y_t)$.

(a) The density of the zero-coupon bond price $P_T^z(S)$ conditional upon z_t , for the Vasicek model ($z = x$) and the CIR model ($z = y$) is given by

$$f_{P_T^z(S) | z_t}^S(u) = \frac{1}{u B^z(\Delta)} f_{z_T | z_t}^S \left(\frac{1}{B^z(\Delta)} \ln \frac{A^z(\Delta)}{u} \right), \quad (20)$$

where $u \geq 0$ for the Vasicek model and $0 \leq u \leq 1$ for the CIR model.

As for the combined model, it reads as

$$f_{P_T^S(S) | \mathcal{F}_t}^S(u) = \int_0^1 \frac{1}{v} f_{P_T^x(S) | x_t}^S(u/v) f_{P_T^y(S) | y_t}^S(v) dv, \quad u \geq 0. \quad (21)$$

(b) The density of the spot rate $L^{z^\phi}(T, S) := F_T^{z^\phi}(T, S)$ conditional upon $\mathcal{F}_t = \mathcal{F}_t^x \vee \mathcal{F}_t^y$,[†] $z \in \{x, y, r\}$, is given by

$$f_{L^{z^\phi}(T, S) | \mathcal{F}_t}^S(u) = \frac{\Delta}{(1 + \Delta u)^2} f_{P_T^z(S) | \mathcal{F}_t}^S \left(\frac{1}{1 + \Delta u} \right). \quad (22)$$

3. Model calibration

Calibrating a short-rate model z to the market amounts to (i) selecting a set of n financial products, called calibration instruments hereafter, and (ii) looking for the set of parameters Ψ^z minimizing the discrepancies between the market- and model-implied prices of the calibration instruments. The

[†] Since x and y are independent, $L^x(T, S) | \mathcal{F}_t$ is equivalent to $L^x(T, S) | x_t$, $L^y(T, S) | \mathcal{F}_t$ is equivalent to $L^y(T, S) | y_t$ and $L^r(T, S) | \mathcal{F}_t$ is equivalent to $L^r(T, S) | (x_t, y_t)$.

time- t no-arbitrage price of a caplet on the spot rate $L(T, S) = F_T(T, S)$, with strike K , fixing date T and payment date $S > T$ can be obtained from the $\mathbb{Q}^S$ density of $L(T, S)$:

$$\begin{aligned} \text{Cpl}_t(T, S, K) &= M_t \mathbb{E} \left[\frac{(L(T, S) - K)^+ \Delta}{M_S} \middle| \mathcal{F}_t \right] \\ &= P_t(S) \mathbb{E}^S \left[(L(T, S) - K)^+ \Delta \middle| \mathcal{F}_t \right] \\ &= \Delta P_t(S) \int_K^{+\infty} (v - K) f_{L(T, S)|\mathcal{F}_t}^S(v) dv, \quad (23) \end{aligned}$$

where $P_t(S)$ is the price of the zero-coupon bond with maturity S at time t and $f_{L(T, S)}^S$ is the density of the spot rate $L(T, S)$ provided in equation (22). The model prices are obtained by evaluating the right-hand side of equation (23), replacing $f_{L(T, S)}^S$ by $\hat{f}_{L(T, S)}^S$ in proposition 3 for each of the considered models. Calibration based on the minimization of the mean squared error of implied volatilities is often favoured in practice and we cover this approach in section 3.1. In section 3.2, we introduce a novel calibration technique whose aim is to minimize the distance between the model-implied density, $\hat{f}_{L(T, S)}^S$, and the market-implied one, $f_{L(T, S)}^S$.

3.1. Calibration of the implied-volatility curve

The caplet pricing equation provides us with the formula to compute the caplet prices under the three considered models by plugging in the density function specified in equation (22). Under the assumption of normal volatility and the Bachelier formula, we can obtain the implied volatility associated with caplet products. Based on the implied volatility, the considered models are calibrated to the market caplet volatility surface, and the fitting performance is measured in terms of the mean squared error (MSE) of implied volatilities. We thus calibrate each of the considered models by looking for the parameters Ψ^z which minimize the MSE between the model-implied and market-implied volatilities:

$$\Psi^{z*} := \arg \min_{\Psi^z} \frac{1}{n} \sum_{i=1}^n (\sigma_i^M - \sigma_i^z)^2 \quad (24)$$

where σ_i^M and σ_i^z are the market- and model-implied volatilities for caplet i , respectively.

3.2. Calibration of the forward rate density

The caplet pricing equation allows us to infer the implied forward rate density from market data in a model-free way.[†] Indeed, differentiating twice the expression for the caplet price in equation (23) with respect to the strike K , we obtain

$$\hat{f}_{L(T, S)|\mathcal{F}_t}^S(K) = \frac{\frac{\partial^2}{\partial K^2} \text{Cpl}_t(T, S, K)}{\Delta P_t(S)}. \quad (25)$$

To get an accurate estimation of the density based on equation (25), we need to obtain the market discount curve

[†] See Breeden and Litzenberger (1978) for a seminal contribution and Brigo and Mercurio (2007) for a textbook treatment in the context of caplet pricing.

and the price quotations for caplets on a fine grid of strikes. Since only a limited number of products are actively traded and quoted on the fixed-income market, we have to resort to interpolation methods. Special care needs to be taken to obtain a valid (i.e. non-negative) and smooth density.

To this end, we use the arbitrage-free smoothing technique of the implied volatility surface developed by Fengler (2009) which we apply to the caplet prices. The procedure takes the quoted caps data and obtains the caplet prices via caplet stripping. Then, a cubic spline smoothing technique is applied to caplet prices, augmented with appropriately chosen linear constraints to enforce no-arbitrage restrictions. By eliminating arbitrage opportunities, this approach precludes the occurrence of negative local volatilities or negative transition probabilities and thus ensures that the forward rate density is well specified.[‡] The method proposed by Fengler (2009) is closely related to the literature on non-parametric estimation of risk-neutral transition densities under shape constraints (Ait-Sahalia and Duarte 2003, Yatchew and Härdle 2006). We decided to adopt it instead of other competing approaches due to its ease of implementation; its good theoretical properties—inherited from the theory on natural smoothing splines under suitable shape constraints; and its applicability in contexts with scarce calibration data. See also Tavin (2012) for alternative approaches to the smoothing of the implied-volatility surface which do not preclude the presence of arbitrage opportunities and impose stronger parametric restrictions on the modelling.

To calibrate the model-implied density to match the market-implied one, we need a criterion assessing how far the model-implied density is from the reference one. A common approach taken in the literature relies on a moment-based approach which focuses on a few key moments of the distribution (Guillaume and Schoutens 2013). This problem has also been tackled by adopting an information-theoretic point of view, in which all moments are taken into account. In particular, Cont and Tankov (2004) introduce a regularization term based on relative entropy in a calibration problem based on MAE minimization to enforce uniqueness and stability of the solution. See also Brody and Hughston (2002) for an application of the Shannon entropy to term structure modelling. We build on this existing literature and select the measure based on Kullback–Leibler (KL) divergence as our calibration criteria.[§] Specifically, the KL divergence between two densities (p, q) is the relative entropy, and is based on (differential) Shannon's entropy:

$$\langle p|q \rangle := \int p(x) \log \left(\frac{p(x)}{q(x)} \right) dx, \quad (26)$$

[‡] A detailed outline of this procedure is provided in appendix 2.

[§] The KL divergence is not a distance because the triangular inequality may fail to hold. See Basseville (2013) for a review of statistical applications using divergence measures.

[¶] We opt for the KL divergence compared to the L^2 metric for density calibration as this measure does not properly scale with the distribution—exacerbating (resp. minimizing) discrepancies between densities in high (resp. low) density regions—which would impact calibration accuracy. Other choices, such as e.g. the Hellinger distance used in Tavin (2012), could be entertained and are left for future research.

where p is assumed to be absolutely continuous with respect to q .

Note that, given the asymmetry of the KL divergence, $\langle p|q \rangle$ will typically be different from $\langle q|p \rangle$. As a consequence, basing the optimization program on only one of them would result in different outcomes. Additionally, optimization based on the KL divergence might face some challenges in practice as the KL divergence is not defined when p is not absolutely continuous with respect to q . In our application, we circumvent these issues by considering the Jensen–Shannon (JS) divergence which is a symmetrized and smoothed version of the KL divergence. Specifically, the JS divergence between two densities (p, q) is defined as

$$\text{JSD}(p|q) = \frac{1}{2}\langle p|m \rangle + \frac{1}{2}\langle q|m \rangle, \quad (27)$$

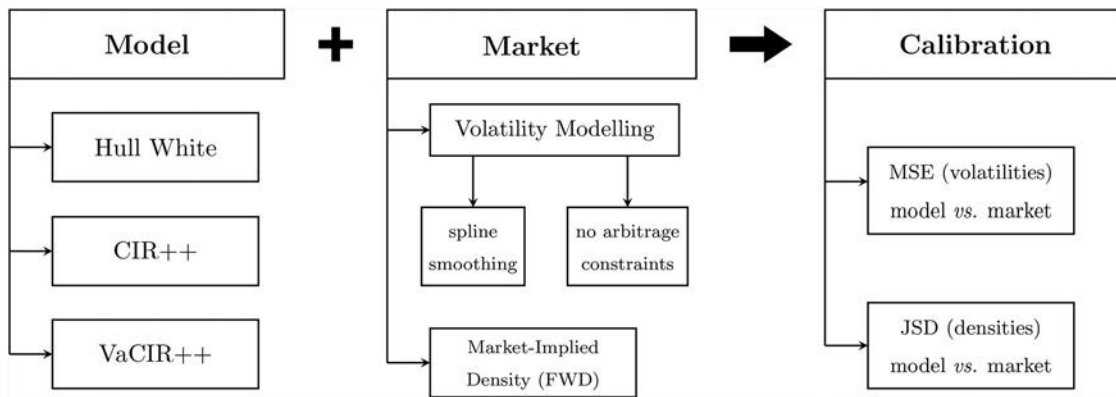
where $m = \frac{1}{2}(p + q)$. By comparing both densities to their average, the JS divergence is always ensured to be bounded as the denominator in the log is always nonzero whenever the numerator is nonzero. Therefore, the JS divergence can be used to compare densities with different supports. Additionally, the measure is bounded at zero and reaches this lower bound only when $p = q$ almost everywhere.

Equipped with a measure accurately quantifying the goodness of fit between densities, we can calibrate each of the considered models by choosing as parameters the set Ψ^z which minimize the JS divergence between the model-implied and the market-implied densities

$$\Psi^{z^*} := \arg \min_{\Psi^z} \text{JSD}(f_{L(T,S)}^S | f_{L^z(T,S)}^S) \quad (28)$$

where $f_{L(T,S)}^S$ and $f_{L^z(T,S)}^S$ are respectively the market- and model-implied forward rate densities.

3.3. Summary of calibration algorithm



As summarized in the diagram 3.3, we use algorithm 1 and algorithm 2 to calibrate the model over a discrete grid of strikes $K \in \{K_0, K_1, K_2 \dots K_{N-1}, K_N\}$ for certain lower and upper bounds of K_0 and K_N .[†]

[†] Volatility modelling and market-implied density are only needed under the calibration method using JS divergence. The calibration method minimizing the mean squared error is based on the implied volatility data directly obtained from caplet stripping.

Algorithm 1 Calibration of the implied (normal) volatility curve

- 1: Calculate the caplet prices under the chosen model (Hull–White, CIR++, or VaCIR++) according to equation (23).
- 2: Obtain the market-implied (normal) volatilities for the corresponding caplet products.
- 3: Use mean squared error as the target criterion to calibrate the three models by matching the market volatility curve according to equation (24).

Algorithm 2 Calibration of the forward rate density

- 1: Calculate the model-implied density of the forward rate under the chosen model (Hull–White, CIR++ or VaCIR++) according to proposition 2 and proposition 3.
- 2: Compute the market-implied density for the forward rate.
 - Retrieve caplet prices from market cap data via caplet stripping
 - Implement arbitrage-free smoothing technique of volatility according to the algorithm presented in appendix 2
 - Obtain the market-implied density over a fine grid according to equation (25).
- 3: Use JS divergence as the target criterion to calibrate the three models by matching the market-implied density according to equation (28).

3.4. Multi-period extension

The calibration approach outlined in section 3.2 only considers caplet products with a single pair of fixing and payment dates (T, S) for different strikes $K \in \{K_0, K_1, \dots, K_N\}$ whose underlyings correspond to the same spot rate $L(T, S)$. One might also be interested in pricing multiple caplet products with different fixing and

payment dates using a common short-rate model. Although several routes are possible, this can be easily achieved with the following two-step procedure. First, we retrieve caplet prices for products with pairs of fixing and payment dates denoted by (T^i, S^i) , $i = 1, \dots, M$, for different strikes $K^i \in \{K_0^i, K_1^i, \dots, K_{N_i}^i\}$ from market cap data using caplet stripping and we construct the caplet price curve for a finer grid of strikes via the arbitrage-free smoothing technique of the implied-volatility surface (Fengler 2009). Second, we can

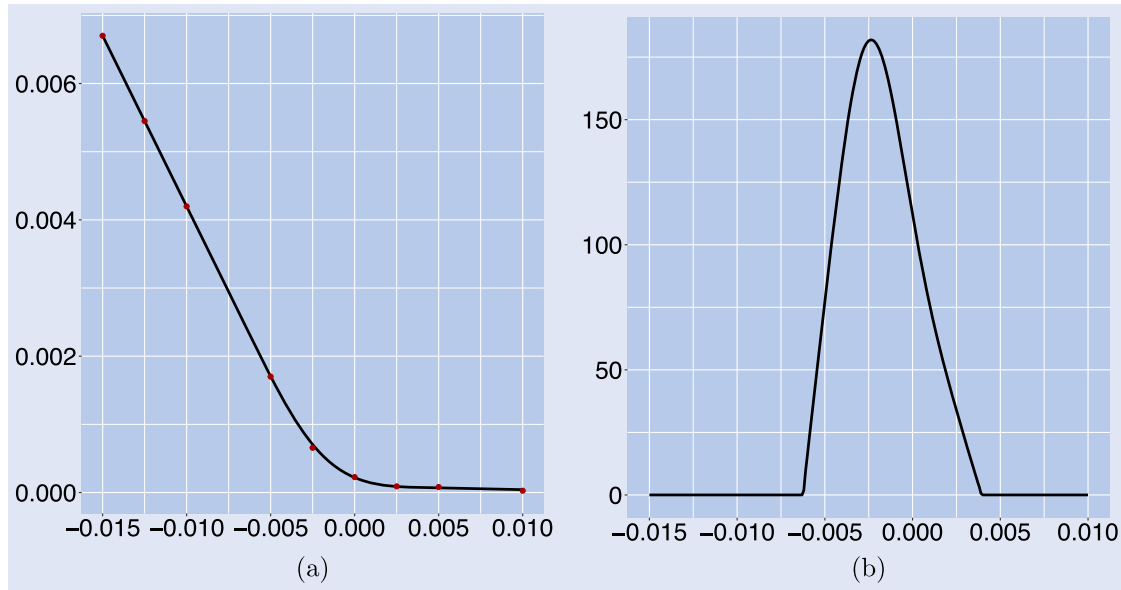


Figure 1. Interpolated caplet prices and market-implied probability density under $\mathbb{Q}^S$ constructed based on the cap market data (quoted in normal volatility) on March 31, 2016. Panel (a) gives the market caplet prices (dots) computed by the Bachelier formula alongside with interpolated price curve obtained by the no-arbitrage cubic spline smoothing technique. Panel (b) provides the $L(6M, 1Y)$ forward rate density implied from the interpolated caplet prices using equation (25). (a) Interpolated price curve and (b) implied density of $L(6M, 1Y)$.

calibrate each of the considered short-rate models by choosing the set of parameters Ψ^z which minimizes the sum of all JS divergence

$$\Psi^{z*} := \arg \min_{\Psi^z} \sum_{i=1}^M JSD(f_{L(T^i, S^i)}^{S^i} | f_{L^z(T^i, S^i)}^{S^i}) \quad (29)$$

where $JSD(f_{L(T^i, S^i)}^{S^i} | f_{L^z(T^i, S^i)}^{S^i})$ is the JS divergence value between the market- and model-implied densities for the corresponding spot rates $L(T^i, S^i)$. A weighting scheme could further be applied to the elements of the sum in equation (29) to reflect, e.g. relative differences in liquidity across fixing and payment dates.

4. Empirical application

In this section, we present an empirical application where we conduct a thorough performance comparison of the three competing models (namely, Hull–White, CIR++ and VaCIR++) under the two calibration schemes introduced in section 3. More specifically, we take the 6-month forward rate as an example to evaluate the pricing accuracy and the quality of fitting the market-implied-volatility curve. We further investigate the distribution of the forward rate in 6 months and illustrate the model performance in terms of density matching under the three considered frameworks in the negative interest rate environment. The rest of this section is organized as follows: section 4.1 presents the data description and sample coverage used for our empirical application and illustrates the non-arbitrage smoothing technique implemented to construct the market-implied density curve; Section 4.2 covers the calibration results based on the minimization of implied volatility errors; Section 4.3 covers the calibration results based on the minimization of JS divergence where the modelled forward density is matched to the market-implied one.

4.1. Data description and volatility curve smoothing

We use historical data for the 6-month EURIBOR forward rate $F_0(6M, 1Y)$ and the interest rate derivatives whose underlying rate is the 6-month forward rate in 6 months, i.e. $L(6M, 1Y) = F_{6M}(6M, 1Y)$.[†] Observe that the market provides quotations for the implied volatilities of caps whereas, to compute the market-implied density for the forward rate $L(6M, 1Y)$, we need caplet prices associated with the latter.

Therefore, we start by collecting monthly normal volatility quotes from Refinitiv (formerly known as Eikon, Thomson Reuters) for the 6-month EURIBOR caps quoted over a grid composed of 13 different strikes. The sample period spans from March 31, 2016, to January 29, 2021, during which the 6-month EURIBOR and 6-month forward rate are in the negative territory. Then we use the Bachelier formula to get the cap prices and apply the stripping approach proposed in Hagan and Konikov (2004) to extract caplet prices.

The market-implied volatilities for caplets are obtained by inverting the Bachelier formula. The implied forward rate density can be computed from equation (25), which involves the numerical estimation of the second-order derivatives. Thus caplet prices for a given fine grid of strikes are required to guarantee an accurate approximation. As explained in section 3.2, we follow the procedure introduced in Fengler (2009) to ensure a well-specified forward rate density.

In figure 1, we illustrate the performance of the smoothing approach from two perspectives, the fit of market prices and the behaviour of the probability density. Figure 1(a) presents the interpolated prices and the market quoted prices for the caplet products of interest, $Cpl(0, 6M, 1Y)$. It can be observed that all market quotations are in line with the

[†] The discount curve and all the related interest derivatives are based on the 6-month EURIBOR rates for consistency.

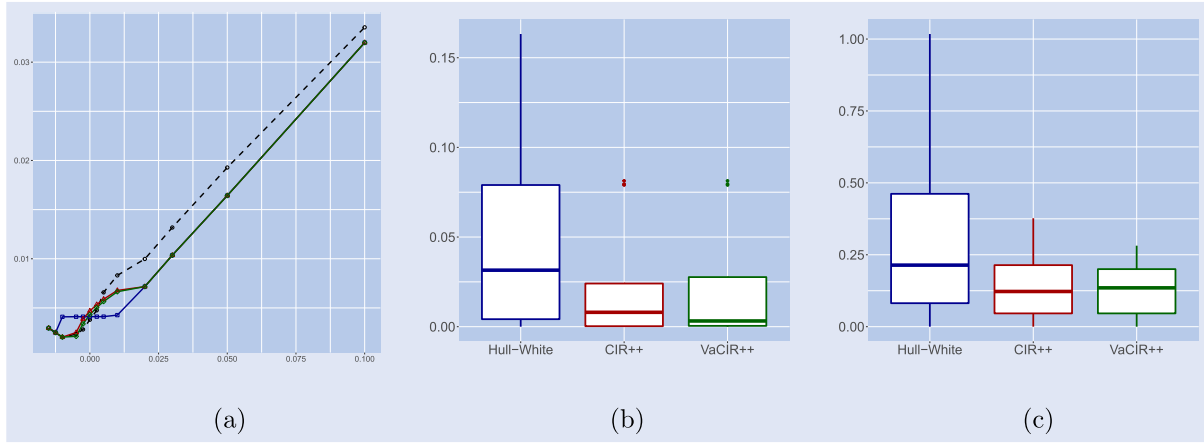


Figure 2. Implied-volatility curves with associated squared errors (in bps) and relative volatility errors on March 31, 2016, where the models are calibrated on the criteria of implied volatility error minimization. Panel (a) shows the implied-volatility curves with respect to the strike for the market (dashed), Hull–White ($\square$), CIR++ ($\triangle$) and VaCIR++ ($\diamond$) models. Panel (b) and Panel (c) provide respectively the box plots of the associated squared and relative errors for the three candidate models. (a) Implied-volatility curve; (b) implied vol squared error and (c) implied vol relative error.

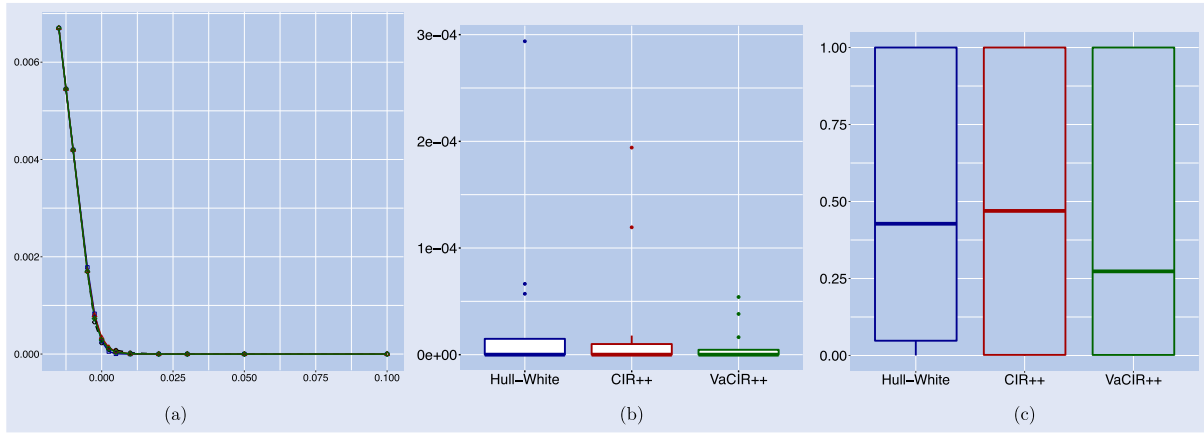


Figure 3. Caplet price curves with corresponding squared errors (in bps) and relative pricing errors on March 31, 2016, where models are calibrated on the criteria of implied volatility error minimization. Panel (a) displays the market and modelled caplet prices with respect to the strike for the market (dashed), Hull–White ($\square$), CIR++ ($\triangle$) and VaCIR++ ($\diamond$) models. Panel (b) and panel (c) provide respectively the box plots of the corresponding squared errors and relative errors for the three candidate models. (a) Caplet price curve, (b) price squared error and (c) price relative error.

smoothed price curve, indicating that the interpolation technique is performing well as an accurate caplet valuation tool. Figure 1(b) shows that the market-implied density is non-negative and is positive in the range from approximately -0.7% to 0.3% . This indicates that the 6-month forward rate in 6 months is distributed in a quite narrow range and the probability of the potential value is the highest around -0.22% which is consistent with the level of the current forward rate. This illustrates the property that the expectation under the corresponding forward measure of the future forward rate matches the current forward rate.

4.2. Calibration results for the implied-volatility curve

Equation (23) provides the model-implied caplet prices. The corresponding caplet volatilities, obtained via the Bachelier formula, can thus be compared with those extracted from market quotes as explained in section 3.1. The model is then calibrated according to equation (24). We provide the

calibration performance on the first available date from our sample as an illustration.

We plot in figure 2 the modelled implied volatilities, with associated squared errors in basis points and relative errors, using the parameters calibrated based on fitting the market-implied volatilities. Figure 2(a) shows the implied-volatility curves for each of the three calibrated models as well as the implied volatilities extracted from stripped market quotes, which serve here as calibration target. We observe that the CIR++ and the VaCIR++ models capture quite well the in-/at-the-money caplet volatilities but fail to capture accurately the data of the deep out-of-the-money products. The Hull–White model provides the worst fit as the modelled curve deviates from the market one for all products. Furthermore, the box plots of the associated squared and relative errors, displayed in figures 2(b,c) respectively, provide further details on the relative performance of the three models considered. The Hull–White model has the largest median value and dispersion for both the squared and relative errors while the error distribution of the VaCIR++ model

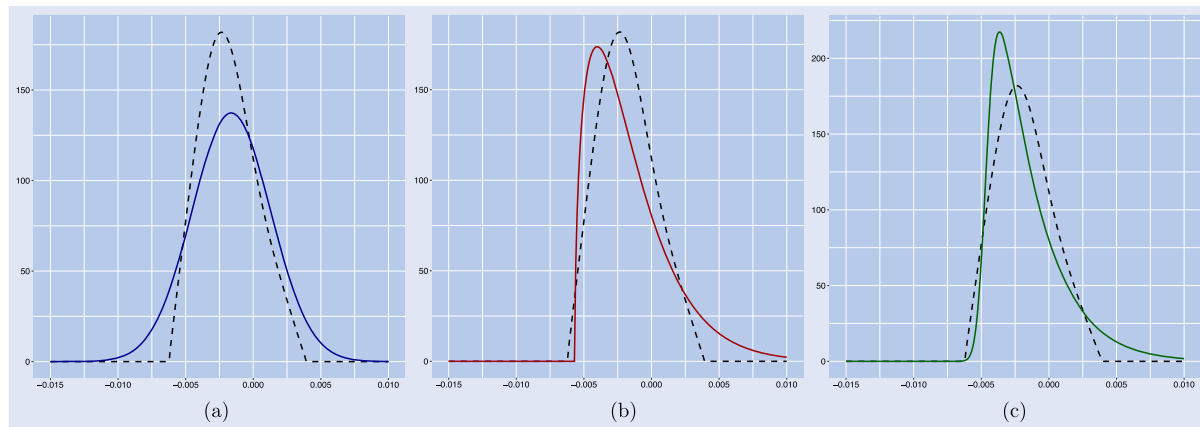


Figure 4. Density curves of the forward rate $L(6M, 1Y)$ on March 31, 2016, where models are calibrated on the criteria of implied volatility error minimization. Panel (a), panel (b) and panel (c) exhibit the market-implied forward density curve and the modelled forward density curves (dashed lines) and under the Hull–White, the CIR++ and VaCIR++ models, respectively (solid lines). (a) Hull–White model, (b) CIR++ model and (c) VaCIR++ model.

is close to the CIR++ one but exhibits a slightly lower median value (squared errors) and less dispersion (relative errors).

Figure 3 shows the same results as figure 2 but expressed in terms of prices rather than volatilities. It can be observed that all three considered models provide a close fit to the market price curve. When comparing between models, the inferior performance of the Hull–White model is further confirmed by pricing accuracy. In particular, the Hull–White model has the highest value in terms of squared errors and comparable relative errors to the CIR++ model. The VaCIR++ model provides the best performance in terms of both the squared and relative pricing errors. Not surprisingly, figure 3(a) indicates that a small implied volatility error for the in/at the money caplets can lead to a sizable discrepancy in prices while the out of the money caplets are subject to large differences in volatility but have almost indistinguishable price curves. This observation highlights the drawbacks of the calibration approaches fitting the implied volatilities or caplet prices. More specifically, the far out of the money options tend to have extremely low prices but rather high implied volatilities. Therefore, putting equal weight on the error of those products in the calibration target could sacrifice the accuracy in modelling the exact prices and volatilities for the at/in the money products. Thus a weighting scheme should be incorporated such that the information in such options would be substantially discounted. However, this approach would still be essentially focusing on a subset of products and would likely result in a poor overall fit.

We further investigate the distribution of the forward rate under the three considered models compared to the market-implied one. Figure 4 displays respectively the density curves for the Hull–White, the CIR++ and the VaCIR++ models against the market-implied one. We observe that in general, the three models fail to match the market-implied forward rate density especially around the centre of the distribution and in the right tail. In particular, it can be observed from figure 4(a) that the forward rate distribution under the Hull–White model is symmetric and has a lower peak than the market-implied one. As shown in figures 4(b,c), the density under the CIR++ and VaCIR++ models provides a close fit

on the left-hand side of the distribution but not in the centre or in the right tail. This observation suggests that the calibration approach focusing on a few implied volatility data points does not guarantee a match of the entire forward rate distribution.

4.3. Calibration results for the forward rate density

Equipped with the smoothed caplet price curve computed over a fine grid of strikes, we now turn to the estimation of the market-implied forward rate distribution according to equation (25). We then calibrate the three considered models according to algorithm 2 to closely match the market-implied density by minimizing the discrepancy (measured by JS divergence) between the model-implied and market-implied curves. As in the previous section, we take the first available date from our sample and investigate the distribution of the forward rate under the three candidates compared to the market-implied one.

We plot in figure 5 the density curves obtained by calibrating each of the three models (solid lines). Figure 5(a) shows that the calibrated density under the Hull–White model has a right-shifted distribution and is more symmetric, compared to the market-implied one. Furthermore, it provides a closer fit to the market-implied density on the right-hand side but deviates significantly on the left-hand side. In addition, we plot in figure 5(b) the density under the CIR++ model, we observe that it has a slightly higher peak than the market-implied density but it captures quite well the skewness exhibited by the market density. In particular, it provides an almost perfect fit on the left-hand side although it fails to match the right-hand tail in the forward rate distribution. Lastly, the calibrated density obtained from the VaCIR++ model is displayed in figure 5(c), and we conclude that the combined model matches well both tails and captures the overall market-implied density curve better than the Hull–White and CIR++ models. Furthermore, the superior performance of the VaCIR++ is confirmed by the significantly lower JS divergence value on this calibration date compared to the alternative models (see figure 9). This observation is consistent with the argument that

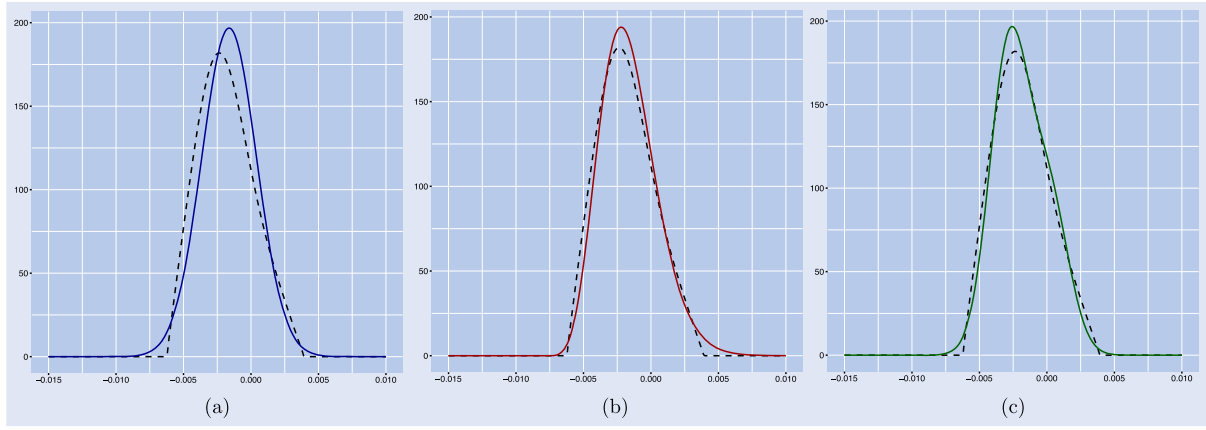


Figure 5. Density curves of the forward rate $L(6M, 1Y)$ on March 31, 2016, where models are calibrated on the criteria of JS divergence minimization. Panel (a), panel (b) and panel (c) exhibit the market-implied forward density curve (dashed lines) and the modelled forward density curves under the Hull–White, CIR++ and VaCIR++ models, respectively. (a) Hull–White model, (b) CIR++ model and (c) VaCIR++ model.

the VaCIR++ model will perform at least as well as its underlying component processes and it will be inclined to place a higher weight on the component process exhibiting better calibration performance, the CIR++ process in this case. It is worth noting that the model-implied densities obtained from the density matching calibration approach provide a tighter fit to the market-implied one compared to the results displayed in figure 4 which are based on the implied volatility MSE criterion.

We further illustrate in figure 6 the pricing accuracy of the three models using the calibrated parameters based on density matching. Figure 6(a) compares the model-implied caplet prices to the market data.[†] It can be concluded that all three models are able to provide a quite accurate valuation for the caplet products. The squared and relative pricing errors under the three models are plotted in figures 6(b,c). The bar in the boxplots represents the median of the errors while the lower and upper bounds of each box represent the first and third quartiles respectively. We observe that the VaCIR++ model provides the lowest pricing errors while the Hull–White and CIR++ models have comparable pricing accuracy. Additionally, we note that high absolute pricing accuracy does not necessarily translate to better performance in terms of relative pricing errors as close to zero prices might lead to high relative pricing errors. Furthermore, high pricing accuracy indicates that the better density fit obtained when adopting the JS divergence calibration technique instead of the price or implied volatility MSE criterion does not come at the cost of larger pricing errors.

In addition, we assess in figure 7 the performance of the three models in terms of how close their implied-volatility curves, arising from minimizing the JS divergence, are to the market-based implied-volatility curve. Figure 7(a) shows that

there are visible discrepancies between the implied volatilities under the CIR++ model and the market ones for caplet products with strikes between 0% and 1%. Outside that range, the VaCIR++ model has superior matching performance of the implied-volatility curve. Furthermore, we note that the calibration focused on density matching also provides convincing fitting performance in terms of implied volatility errors. Therefore, we conclude that the density matching approach provides a more robust way of calibrating interest rate models to fixed income derivative data since it has superior performance in capturing the entire forward distribution, high pricing accuracy and comparable fitting quality in terms of matching implied-volatility curves.[‡]

Thus far, our analysis focused on benchmarking our two-factor VaCIR++ model with the two components, namely the one-factor Hull–White and CIR++ models. Therefore, we also include additional two-factor models in the comparison to capture the added flexibility associated with the use of a two-factor model. We opt for the G2++ model which is a two-factor Gaussian model introducing correlation between its two Brownian motions through an additional parameter ρ . Additionally, we consider the CIR2++ model whose underlying factors are two independent CIR processes.^{§¶} We plot in figure 8 the density curves obtained by calibrating the G2++ and CIR2++ models (solid lines) to the market data. Figure 8(a) shows that the calibrated density under the G2++ model is almost identical to the density obtained from a Hull–White model presented in figure 5(a), namely

[‡] Note that, although our approach takes into account the entire forward rate distribution in the calibration, it does not provide at the moment an explicit way to assign different weights on the calibration products to reflect their different degrees of data reliability (moneyness or liquidity). This is indeed an interesting question, which would increase the practical relevance of our approach. However, we propose to keep this extension for future work. We thank one of the anonymous referees for pointing this out.

[§] The interested reader is referred to chapter 4 of Brigo and Mercurio (2007) for more details on the G2++ and CIR2++ models.

[¶] In terms of model complexity, the one-factor Hull–White and CIR++ models have each four parameters while the G2++ model is parameterized with six parameters. The CIR2++ and VaCIR++ models have eight parameters to estimate.

[†] We note that the market prices are recovered by computing the risk-neutral expectation of the discounted payoffs using the market-implied density as the resulting prices are more comparable with the model-implied ones. This is due to the fact that the prices consistent with the market-implied density, obtained as a by-product of the no-arbitrage volatility curve smoothing, might slightly differ from the actual market prices, which are not necessarily free from arbitrage opportunities.

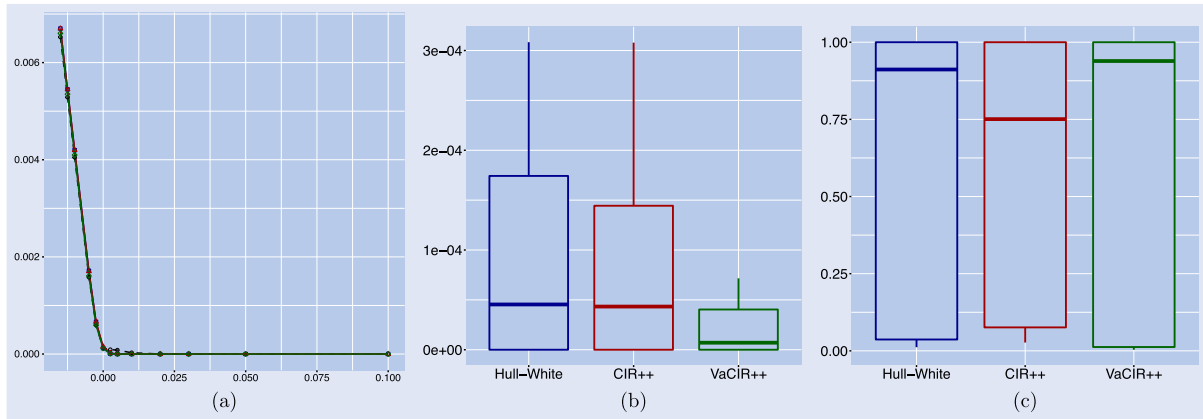


Figure 6. Caplet price curves with corresponding mean squared errors (in bps) and relative pricing errors on March 31, 2016, where models are calibrated on the criteria of JS divergence minimization. Panel (a) displays the market and modelled caplet prices with respect to the strike for the market (dashed), Hull-White ($\square$), CIR++ ($\triangle$) and VaCIR++ ($\diamond$) models. Panel (b) and panel (c) provide respectively the box plots of the corresponding mean squared errors and relative errors for the three candidate models. (a) Caplet price curve, (b) price squared error and (c) price relative error.

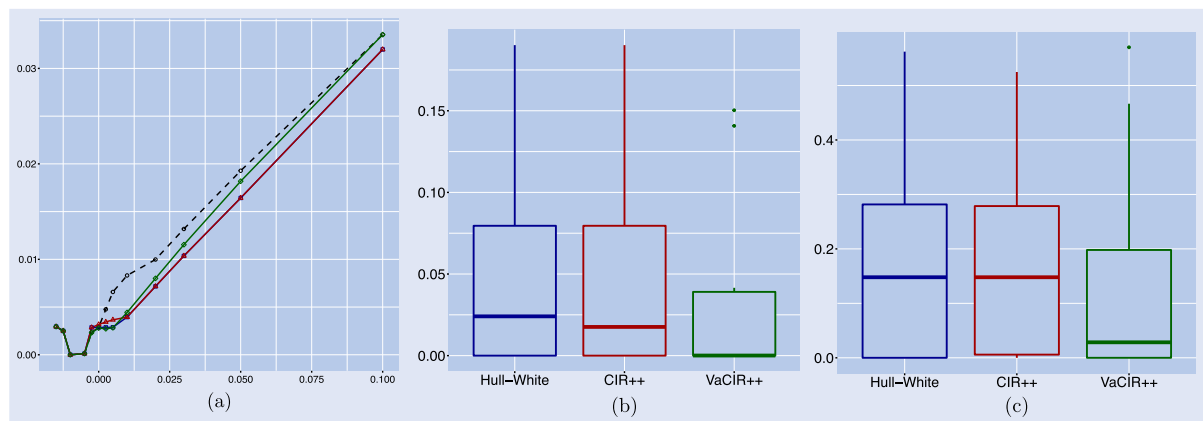


Figure 7. Implied-volatility curves with associated squared and relative volatility errors on March 31, 2016, where the models are calibrated on the criteria of JS divergence minimization. Panel (a) contains the implied-volatility curves with respect to the strike for the market (dashed), Hull-White ($\square$), CIR++ ($\triangle$) and VaCIR++ ($\diamond$) models. Panel (b) and Panel (c) provide respectively the box plots of the associated squared errors and relative errors for the three candidate models. (a) Implied-volatility curve; (b) implied vol squared error and (c) implied vol relative error.

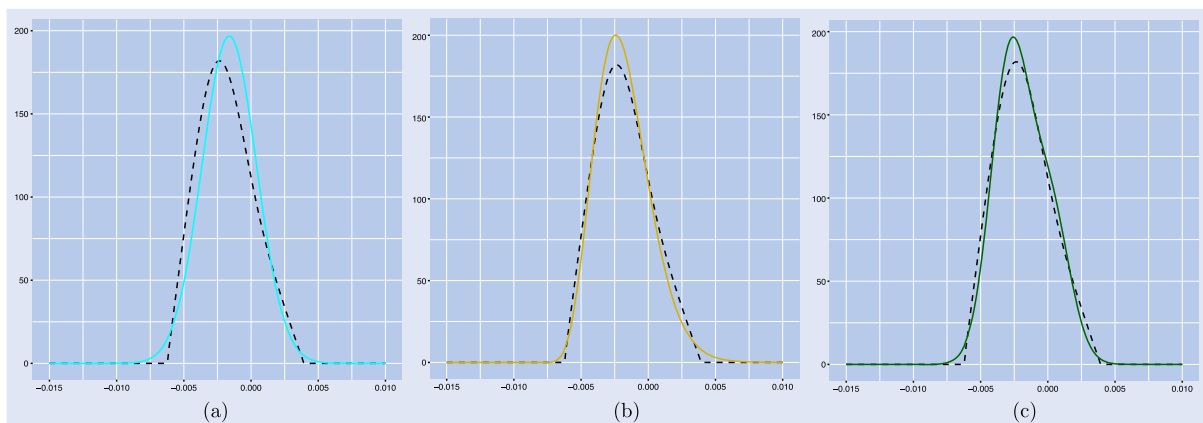


Figure 8. Density curves of the forward rate $L(6M, 1Y)$ on March 31, 2016, where models are calibrated on the criteria of JS divergence minimization. Panel (a), Panel (b) and panel (c) exhibit the market-implied forward density curve and the modelled forward density curves (dashed lines) and under the G2++, CIR2++ and VaCIR++ models, respectively (solid lines). (a) The G2++ model; (b) the CIR2++ model and (c) VaCIR++ model.

it has a right-shifted distribution and is more symmetric, compared to the market-implied one. It is worth noting that the G2++ model provides indistinguishable performance to its one-factor counterpart (Hull–White) which indicates that the additional flexibility from the two-factor model leads to limited added value for our application.[†] In addition, the calibrated density obtained from the CIR2++ model is displayed in figure 8(b), and we conclude that it provides a similar but closer fit to the market-implied density compared to its one-factor counterpart. Comparing both figures to figure 8(c), the VaCIR++ model provides a better overall fit to the market-implied density curve than the two-factor benchmarks. The superior performance of the VaCIR++ is confirmed by the significantly lower JS divergence value on this calibration date (see figure 9).

We further investigate how model performance—measured in terms of matching the market-implied density curve—evolves over time. To this end, the five models considered are calibrated on a monthly basis and we report in figure 9 the JS divergence for each of them. When comparing the VaCIR++ model to its two one-factor component models, the best-in-class property of VaCIR++ is preserved through time. In particular, it always has the lowest JS divergence, indicating closest matching to the market-implied density curve. Also, we can observe that the CIR++ model exhibits better performance than the Hull–White model which makes the CIR++ model an attractive candidate as a single-factor model—with the shift providing some flexibility to deal with negative rates. Additionally, we observe that the performance of the G2++ and Hull–White models in terms of matching the market-implied forward rate density are essentially indistinguishable. See the light blue marks (G2++ model) in figure 9 which are almost perfectly aligned with the dark blue curve (Hull–White model). In fact, the estimated parameter ρ under the G2++ model is rather close to -1 over the whole sample period. In this case, the G2++ model degenerates into a one-factor short-rate process. This observation is consistent with the results of the empirical application presented in chapter 4 of Brigo and Mercurio (2007). Furthermore, the VaCIR++ model outperforms the CIR2++ benchmark on most calibration dates with only a few exceptions where the CIR2++ model has a slightly lower JS divergence. Lastly, we can observe that the added flexibility of the CIR2++ provides superior performance over the one-factor CIR++, in particular over the first part of the sample. To summarize, the VaCIR++ model is more appealing both theoretically—allowing for negative values, skewness, and no strict lower bound—and empirically as it enhances the fitting performance compared to the four benchmark models and improves calibration stability over time.

Based on the density obtained, we characterize the variation of the implied conditional moments of the forward rate $L(T, S)$ (under the S -forward measure) over time.[‡] In this manner, we compute the first four conditional moments

for the 6-month forward rate in 6 months over the whole sample period and investigate how well each of the five competing models can match the market-implied moments (figure 10). It can be observed from figure 10(a) that the five models accurately track the market-implied mean overall—with the Hull–White, CIR++ and G2++ models exhibiting slight discrepancies in the first 2 years of the sample period. Figure 10(b) shows that the Gaussian models (Hull–White and G2++) have the worst performance in terms of tracking the market-implied standard deviation while the CIR2++ model provides the best fit overall. For the skewness, illustrated in figure 10(c), we observe that the conditional skewness of the forward rate distribution under the Hull–White and G2++ models is almost zero indicating that these two models are unable to capture the skewness featured in the data. The conditional skewness under the CIR++ and CIR2++ is positive, though higher than the market-implied one in general. The conditional skewness under the VaCIR++ model is the closest to the market-implied one, except for a few months in 2018 and 2019 where the model performance is comparable to the CIR++ and CIR2++ models. Lastly, figure 10(d) illustrates that the excess kurtosis under the Hull–White and G2++ models is close to zero while the densities implied by the CIR++ and CIR2++ models always exhibit positive excess kurtosis. In both cases, the model-implied excess kurtosis is overall higher than the market-implied one. The excess kurtosis implied by the VaCIR++ model provides, however, a better tracking of the market-implied one, especially in the first part of the sample. The jumps observed in the conditional skewness and excess kurtosis for the VaCIR++ model might be caused by local minima in the optimization and warrant further investigation. We also note that the CIR++ and VaCIR++ models have similar levels of JS divergence during those months but this does not necessarily imply that the two models have similar densities or conditional moments. Indeed, the two model-implied distributions could be equally similar to the market one while being noticeably different from each other.

Finally, focusing on the forward rate density and its conditional moments also provides us with relevant market insights for risk management and hedging applications. As shown in figure 10(a,b), the mean and standard deviation of the future forward rate are stable at the beginning of the sample period. In particular, for the level of the mean, a significant downturn followed by an increasing trend is observed during 2019 while the period from 2020 to 2021 is dominated by a downward shift. The standard deviation, on the other hand, exhibits a significant upward movement followed by a short decline during 2019 and is then strongly trending upward during the period from 2020 to 2021. This more volatile evolution of interest rates during the COVID-19 period is likely due to the increased uncertainty in financial markets and concerns for funding conditions resulting from the crisis. Furthermore, figure 10(c) shows that the conditional skewness is generally positive with a sharp increase at the onset of the COVID-19 crisis followed by a steady decrease over the rest of the sample period. We observe a similar behaviour in figure 10(d) for the excess kurtosis during COVID. Overall, excess kurtosis is negative over our sample, indicating a lower tail-thickness compared to a Gaussian distribution.

[†] Although the two-factor G2++ model has limited contribution in our application compared to the one-factor Hull–White model, it might lead to improved performance in empirical applications where the correlation between forward rates plays an essential role in the calibration; e.g. for swaption contracts.

[‡] The market- and model-implied conditional moments are computed numerically using the corresponding forward rate density.

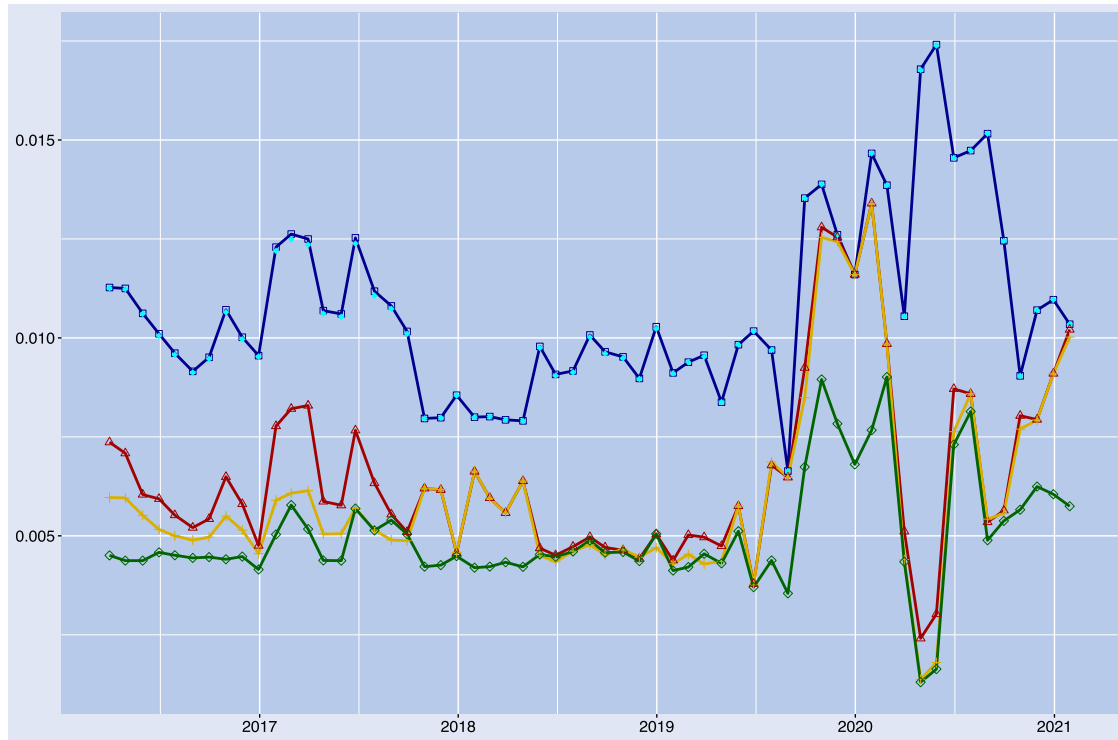


Figure 9. Evolution over time of the JS divergence resulting from the density matching calibration under the Hull-White ($\square$, dark blue), CIR++ (Δ , red), G2++ ($\diamond$, light blue), CIR2++ ($+$, yellow) and VaCIR++ ($\diamond$, green) models.

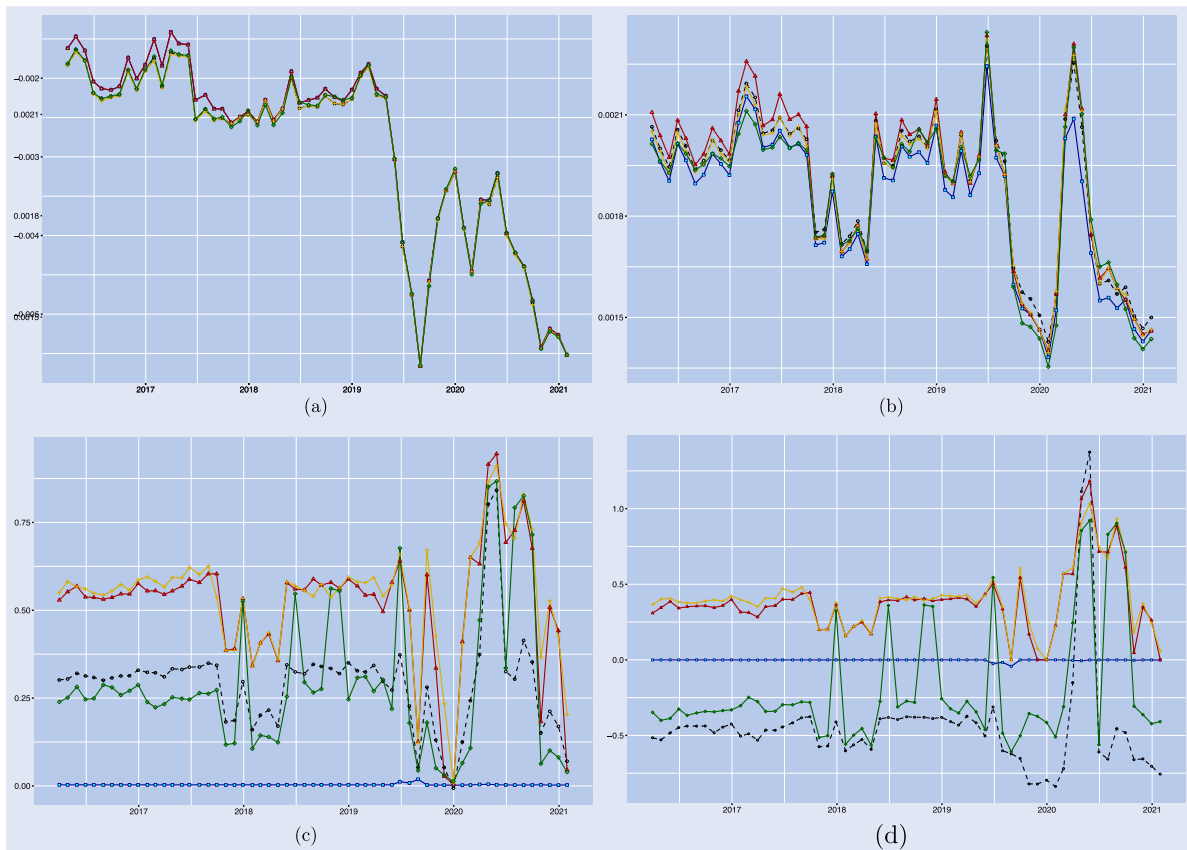


Figure 10. Time series of the $L(6M, 1Y)$ conditional moments for the market ($\circ$) and for the Hull-White ($\square$, dark blue), CIR++ (Δ , red), G2++ ($\diamond$, light blue), CIR2++ ($+$, yellow) and VaCIR++ ($\diamond$, green) models. (a) Conditional mean; (b) conditional standard deviation; (c) conditional skewness and (d) conditional excess kurtosis.

5. Conclusion

We propose a simple asymmetric short-rate model that does not display a strict lower bound and is well suited to capture

the salient features of the negative interest rate environment. Our two-factor model, called the VaCIR++ model, combines the advantages of the Hull-White model (no strict lower bound) and the CIR++ model (positive skewness) while

maintaining analytical tractability. Our framework delivers semi-analytical expressions for the forward rate density and caplet pricing which allows us to tackle the challenges encountered by standard benchmark models for the modelling and pricing of interest rate derivatives in a negative interest rate environment.

In addition, we introduce a new calibration procedure based on density matching which alleviates the drawbacks inherent to standard model calibration procedures based on prices or volatilities mean squared error. Precisely, the model is calibrated such that the model-implied density is the closest to the market-implied one in terms of Jensen–Shannon divergence. While in this context, we match the forward rate densities computed under the forward measure, our approach can be broadly applied to other settings where the interest lies in closely matching a chosen target density.

Finally, we illustrate the benefits of our framework in a financial application featuring a time series of caplets whose prices are retrieved by stripping cap implied volatilities. We provide a comparative study of calibration performance in the period of negative interest rates under two calibration criteria—the volatility mean squared error minimization and our proposed density matching approach based on the minimization of the Jensen–Shannon divergence. We note the outperformance of our model relative to the Hull–White and CIR++ models. Moreover, the calibration procedure based on the Jensen–Shannon divergence significantly enhances the matching of the forward densities while preserving high pricing accuracy and comparable fitting quality for implied-volatility curves.

Our work can be extended in several ways. First, our modelling framework can be extended to consider simultaneously multiple forward rate maturities and contract expiry lengths. This would allow us to explore the benefits of extracting information relevant for model calibration using liquid products and apply this calibration to the pricing of less liquid products within a consistent framework. Second, the strength of the economic recovery following the COVID-19 crisis, combined with the escalation of the conflict between Russia and Ukraine, led to a strong revival of inflationary pressures. As major central banks reacted to these events by raising their policy rates, it would be interesting to evaluate how our proposed model and calibration approach perform in a period where interest rates revert back to higher levels.

Acknowledgments

The authors are grateful to Damiano Brigo and Donatien Hainaut for insightful comments on earlier versions of this manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Belgian Federal Science Policy Office [ARC grant number 18-23/089].

ORCID

Frédéric Vrins  <http://orcid.org/0000-0003-1941-0416>

Linqi Wang  <http://orcid.org/0000-0002-8197-7168>

References

- Acharya, V.V., Eisert, T., Eufinger, C. and Hirsch, C., Whatever it takes: The real effects of unconventional monetary policy. *Rev. Financ. Stud.*, 2019, **32**, 3366–3411.
- Ait-Sahalia, Y. and Duarte, J., Nonparametric option pricing under shape restrictions. *J. Econom.*, 2003, **116**, 9–47.
- Basseville, M., Divergence measures for statistical data processing—An annotated bibliography. *Signal. Processing.*, 2013, **93**, 621–633.
- Bauer, M.D. and Chernov, M., Interest rate skewness and biased beliefs. NBER Working Paper, 2021.
- Breeden, D.T. and Litzenberger, R.H., Prices of state-contingent claims implicit in option prices. *J. Bus.*, 1978, **51**, 621–651.
- Brigo, D. and Mercurio, F., A deterministic-shift extension of analytically-tractable and time-homogeneous short-rate models. *Finance Stoch.*, 2001, **5**, 369–387.
- Brigo, D. and Mercurio, F., *Interest Rate Models: Theory and Practice with Smile, Inflation and Credit*, 2006 (Springer Verlag: Heidelberg, Germany).
- Brody, D.C. and Hughston, L.P., Entropy and information in the interest rate term structure. *Quant. Finance*, 2002, **2**, 70–80.
- Cont, R. and Tankov, P., Nonparametric calibration of jump-diffusion option pricing models. *J. Comput. Finance*, 2004, **7**, 1–49.
- Cox, J.C., Ingersoll Jr, J.E. and Ross, S.A., A theory of the term structure of interest rates. *Econometrica*, 1985, **53**, 385–407.
- Duffie, D. and Kan, R., A yield-factor model of interest rates. *Math. Finance*, 1996, **6**, 379–406.
- Fengler, M.R., Arbitrage-free smoothing of the implied volatility surface. *Quant. Finance*, 2009, **9**, 417–428.
- Filipović, D. and Trolle, A.B., The term structure of interbank risk. *J. Financ. Econ.*, 2013, **109**, 707–733.
- Filipović, D. and Kitapbayev, Y., On the american swaption in the linear-rational framework. *Quant. Finance*, 2018, **18**, 1865–1876.
- Filipović, D., Larsson, M. and Trolle, A.B., Linear-rational term structure models. *J. Finance.*, 2017, **72**, 655–704.
- Guillaume, F. and Schoutens, W., A moment matching market implied calibration. *Quant. Finance*, 2013, **13**, 1359–1373.
- Hagan, P. and Konikov, M., Interest rate volatility cube: Construction and use. Bloomberg technical report No. 62, 2004.
- Heider, F., Saidi, F. and Schepens, G., Life below zero: Bank lending under negative policy rates. *Rev. Financ. Stud.*, 2019, **32**, 3728–3761.
- Hull, J. and White, A., Pricing interest-rate-derivative securities. *Rev. Financ. Stud.*, 1990, **3**, 573–592.
- Li, H. and Zhao, F., Nonparametric estimation of state-price densities implicit in interest rate cap prices. *Rev. Financ. Stud.*, 2009, **22**, 4335–4376.
- Longstaff, F.A., Mithal, S. and Neis, E., Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market. *J. Finance.*, 2005, **60**, 2213–2253.
- Recchioni, M.C. and Sun, Y., An explicitly solvable Heston model with stochastic interest rate. *Eur. J. Oper. Res.*, 2016, **249**, 359–377.
- Recchioni, M.C., Sun, Y. and Tedeschi, G., Can negative interest rates really affect option pricing? Empirical evidence from an explicitly solvable stochastic volatility model. *Quant. Finance*, 2017, **17**, 1257–1275.
- Tavin, B., Implied distribution as a function of the volatility smile. *Bank. Markets Invest.*, 2012, **119**, 31–42.
- Trolle, A.B. and Schwartz, E.S., The swaption cube. *Rev. Financ. Stud.*, 2014, **27**, 2307–2353.
- Vasicek, O., An equilibrium characterization of the term structure. *J. Financ. Econ.*, 1977, **5**, 177–188.

Yatchew, A. and Härdle, W., Nonparametric state price density estimation using constrained least squares and the bootstrap. *J. Econom.*, 2006, **133**, 579–599.

Appendices

Appendix 1. Pricing formula for existing framework

We provide below the A^z and B^z functions in equation (3) for $z \in \{x, y, y^\phi\}$. For conciseness, we omit the subscript in the κ, θ, σ parameters.

For the Vasicek model specified in equation (1), we have

$$A^x(\tau) = \exp \left\{ \left(\theta - \frac{\sigma^2}{2k^2} \right) [B(\tau) - \tau] - \frac{\sigma^2}{4k} B^x(\tau)^2 \right\},$$

$$B^x(\tau) = \frac{1}{k} [1 - e^{-k\tau}].$$

For the CIR model specified in equation (2),

$$A^y(\tau) = \left[\frac{2h \exp\{(k+h)\tau/2\}}{2h + (k+h)(\exp\{h\tau\} - 1)} \right]^{2k\theta/\sigma^2},$$

$$B^y(\tau) = \frac{2(\exp\{h\tau\} - 1)}{2h + (k+h)(\exp\{h\tau\} - 1)}.$$

where $h = \sqrt{k^2 + 2\sigma^2}$.

The shifted CIR model is a special case of a general model proposed in Brigo and Mercurio (2001). It can be shown that

$$A^{y^\phi}(\tau) = e^{-\phi\tau} A^y(\tau), \quad B^{y^\phi}(\tau) = B^y(\tau).$$

Appendix 2. Arbitrage-free smoothing of the volatility curve

Following the notations in Fengler (2009), assume that we observe the caplet prices $\mathbf{y} = (y_1, \dots, y_n)$ at the strikes $a = u_0, \dots, u_{n+1} = b$, and the function g represents the natural cubic spline function. For the value and second derivative representation, we set $g_i = g(u_i)$ and $\gamma_i = g''(u_i)$, for $i = 1, \dots, n$. Furthermore, we define $\mathbf{g} = (g_1, \dots, g_n)^\top$ and $\boldsymbol{\gamma} = (\gamma_2, \dots, \gamma_{n-1})^\top$. By definition, $\gamma_1 = \gamma_n = 0$.

We formulate the sufficient and necessary conditions to ensure a valid cubic spline using the two matrices $\mathbf{Q}$ and $\mathbf{R}$ defined below. Let $h_i = u_{i+1} - u_i$ for $i = 1, \dots, n-1$, and the elements q_{ij} , of matrix $\mathbf{Q}$, for $i = 1, \dots, n$ and $j = 2, \dots, n-1$, are given by

$$q_{j-1,j} = h_{j-1}^{-1}, \quad q_{j,j} = -h_{j-1}^{-1} - h_j^{-1} \quad \text{and} \quad q_{j+1,j} = h_j^{-1},$$

for $j = 2, \dots, n-1$ and $q_{ij} = 0$ for $|i-j| \geq 2$.

The matrix $\mathbf{R}$ is symmetric and its elements r_{ij} , for $i, j = 2, \dots, n-1$, are defined by

$$r_{i,i} = \frac{1}{3} (h_{i-1} + h_i), \quad \text{for } i = 2, \dots, n-1$$

$$r_{i,i+1} = r_{i+1,i} = \frac{1}{6} h_i, \quad \text{for } i = 2, \dots, n-2$$

$$r_{ij} = 0, \quad \text{for } |i-j| \geq 2$$

Furthermore, we formulate the spline smoothing problem as a quadratic minimization program. Define vector $\mathbf{y} = (w_1 y_1, \dots, w_n y_n, 0, \dots, 0)^\top$, where the w_i are strictly positive weights and vector $\mathbf{x} = (\mathbf{g}^\top, \boldsymbol{\gamma}^\top)^\top$. Furthermore, define the matrices, $\mathbf{A} = (\mathbf{Q}, -\mathbf{R}^\top)$ and

$$\mathbf{B} = \begin{pmatrix} \mathbf{W}_n & 0 \\ 0 & \lambda \mathbf{R} \end{pmatrix}$$

where $\mathbf{W}_n = \text{diag}(w_1, \dots, w_n)$. The smoothing algorithm of the volatility curve is summarized below.

- (1) Estimate the volatility curve via an initial interpolation of caplet prices with respect to the moneyness, which is defined as the strikes in excess to the forward rate.
- (2) Obtain the implied-volatility curve using the spline smoothing technique under no-arbitrage constraints by solving the quadratic program formulated as follows:

$$\begin{aligned} \min_{\mathbf{x}} \quad & -\mathbf{y}^\top \mathbf{x} + \frac{1}{2} \mathbf{x}^\top \mathbf{B} \mathbf{x}, \\ \text{subject to} \quad & \mathbf{A}^\top \mathbf{x} = 0, \\ & \gamma_i \geq 0, \\ & \frac{g_2 - g_1}{h_1} - \frac{h_1}{6} \gamma_2 \geq P_t(S)(S - T), \\ & -\frac{g_n - g_{n-1}}{h_{n-1}} - \frac{h_{n-1}}{6} \gamma_{n-1} \geq 0. \end{aligned}$$

Appendix 3. Correlation between two factors

Appendix 3.1. Introduce correlation within the affine framework

Assume short rate is defined as $r_t = X_t^1 + X_t^2$ and the two factors in the model $X_t = (X_t^1, X_t^2)$ are described as the following system of SDEs under the risk-neutral Q -measure:

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \end{pmatrix} = \begin{pmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} dt + \begin{pmatrix} \sigma_{11} & 0 \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \begin{pmatrix} \sqrt{X_t^1} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} dW_t^1 \\ dW_t^2 \end{pmatrix},$$

which is equivalent to

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \end{pmatrix} = \begin{pmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} dt + \begin{pmatrix} \sigma_{11} \sqrt{X_t^1} & 0 \\ \sigma_{21} \sqrt{X_t^1} & \sigma_{22} \end{pmatrix} \begin{pmatrix} dW_t^1 \\ dW_t^2 \end{pmatrix}.$$

We can rewrite it as

$$\begin{aligned} dX_t^1 &= \kappa_1(\theta_1 - X_t^1)dt + \sigma_{11}\sqrt{X_t^1}dW_t^1, \\ dX_t^2 &= \kappa_2(\theta_2 - X_t^2)dt + \sigma_{21}\sqrt{X_t^1}dW_t^1 + \sigma_{22}dW_t^2. \end{aligned}$$

Here, the correlation between the two factors X_t^1, X_t^2 is introduced by σ_{21} . If we set $\sigma_{21} = 0$, this framework reduces to the case where we have two independent processes with X_t^1 being a CIR process and X_t^2 being a Vasicek process. In addition, this belongs to the Affine framework given that

$$\begin{aligned} \Sigma \Sigma' &= \begin{pmatrix} \sigma_{11} \sqrt{X_t^1} & 0 \\ \sigma_{21} \sqrt{X_t^1} & \sigma_{22} \end{pmatrix} \begin{pmatrix} \sigma_{11} \sqrt{X_t^1} & \sigma_{21} \sqrt{X_t^1} \\ 0 & \sigma_{22} \end{pmatrix} \\ &= \begin{pmatrix} \sigma_{11}^2 X_t^1 & \sigma_{11} \sigma_{21} X_t^1 \\ \sigma_{11} \sigma_{21} X_t^1 & \sigma_{21}^2 X_t^1 + \sigma_{22}^2 \end{pmatrix}. \end{aligned}$$

The zero coupon bond price is defined as

$$\begin{aligned} P_t^r(T) &= \exp(B(T-t)'X_t + A(T-t)) \\ &= \exp\left(B^1(T-t)X_t^1 + B^2(T-t)X_t^2 + A(T-t)\right). \end{aligned}$$

For the factor loadings in the zero-coupon bond prices, A and B are the solutions to

$$\frac{dB^1(t, T)}{dt} = 1 + \kappa_2 B^1(t, T) - \frac{1}{2} \left(\sigma_{21} B^2(t, T) + \sigma_{11} B^1(t, T) \right)^2,$$

$$\begin{aligned}\frac{dB^2(t, T)}{dt} &= 1 + \kappa_1 B^2(t, T), \\ \frac{dA(t, T)}{dt} &= -\kappa_1 \theta_1 B^1(t, T) - \kappa_2 \theta_2 B^2(t, T) - \frac{1}{2} \sigma_{22}^2 B^2(t, T)^2.\end{aligned}$$

The solutions to $B(t, T)$ are formulated as

$$\begin{aligned}B^1(t, T) &= B^1(T - t) = -\frac{1}{\kappa_1} \left[1 - e^{-\kappa_1(T-t)} \right], \\ B^2(t, T) &= B^2(T - t) \\ &= -\frac{(2 - \sigma_{21}^2 B^1(t, T)^2)(\exp\{(T - t)h\} - 1)}{2h + (h + \kappa_2 - \sigma_{21}\sigma_2 B^1(t, T))(\exp\{(T - t)h\} - 1)}, \\ h &= \sqrt{2\sigma_{21}\kappa_2\sigma_2 B^1(T - t) - \kappa_2^2 - 2\sigma_2^2}.\end{aligned}$$

and the solution to $A(t, T)$ can be obtained by integration.

Appendix 3.2. Introduce correlation by correlated Brownian motions

Assume W_t^1, W_t^2 are independent Brownian motions, and the short rate $r_t = X_t^1 + X_t^2$.

$$\begin{aligned}\begin{pmatrix} dX_t^1 \\ dX_t^2 \end{pmatrix} &= \begin{pmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{pmatrix} \left[\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} \right] dt \\ &\quad + \begin{pmatrix} \sigma_{11}\sqrt{X_t^1} & 0 \\ \rho\sigma_{22} & \sqrt{1 - \rho^2}\sigma_{22} \end{pmatrix} \begin{pmatrix} dW_t^1 \\ dW_t^2 \end{pmatrix},\end{aligned}$$

We can rewrite it as

$$\begin{aligned}dX_t^1 &= \kappa_1(\theta_1 - X_t^1)dt + \sigma_{11}\sqrt{X_t^1}dW_t^1, \\ dX_t^2 &= \kappa_2(\theta_2 - X_t^2)dt + \rho\sigma_{22}dW_t^1 + \sqrt{1 - \rho^2}\sigma_{22}dW_t^2 \\ &= \kappa_2(\theta_2 - X_t^2)dt + \sigma_{22}dW_t^3,\end{aligned}$$

where W_3 is a Brownian motion satisfying $\text{corr}(dW_t^1, dW_t^3) = \rho dt$.

Here, ρ introduce a correlation between the CIR process and the Vasicek process by correlated Brownian motions. If we set $\rho = 0$, this framework reduces to the case where we have two independent processes with X_t^1 being a CIR process and X_t^2 being a Vasicek process. When $\rho \neq 0$, this model is no longer affine since

$$\begin{aligned}\Sigma \Sigma' &= \begin{pmatrix} \sigma_{11}\sqrt{X_t^1} & 0 \\ \rho\sigma_{22} & \sqrt{1 - \rho^2}\sigma_{22} \end{pmatrix} \begin{pmatrix} \sigma_{11}\sqrt{X_t^1} & \rho\sigma_{22} \\ 0 & \sqrt{1 - \rho^2}\sigma_{22} \end{pmatrix} \\ &= \begin{pmatrix} \sigma_{11}^2 X_t^1 & \rho\sigma_{11}\sigma_{22}\sqrt{X_t^1} \\ \rho\sigma_{11}\sigma_{22}\sqrt{X_t^1} & \sigma_{22}^2 \end{pmatrix},\end{aligned}$$

Appendix 4. Proof of proposition 1

Point (a) is obtained by combining equation (11) with equation (12). For point (b), we have the dynamics of the zero-coupon bond price for $z \in \{x, y\}$

$$\frac{dP_t^z(T)}{P_t^z(T)} = z_t dt + B^z(T - t)\sigma_t^z dW_t^z, \quad 0 \leq t \leq T.$$

Applying Ito's quotient rule, we have for $0 \leq t \leq T < S$,

$$\begin{aligned}\frac{d(P_t^z(T)/P_t^z(S))}{P_t^z(T)/P_t^z(S)} &= \frac{dP_t^z(T)}{P_t^z(T)} - \frac{dP_t^z(S)}{P_t^z(S)} - \frac{dP_t^z(T)}{P_t^z(T)} \frac{dP_t^z(S)}{P_t^z(S)} + \left(\frac{dP_t^z(S)}{P_t^z(S)} \right)^2 \\ &= \sigma_t^z (B^z(T - t) - B^z(S - t)) (dW_t^z - \sigma_t^z B^z(S - t)dt).\end{aligned}$$

Combining this expression with (12), the dynamics of the forward rate become

$$\begin{aligned}dF_t^z(T, S) &= \frac{1}{\Delta} d(P_t^z(T)/P_t^z(S)) \\ &= \left(\frac{1}{\Delta} + F_t^z(T, S) \right) \sigma_t^z ((B^z(S - t) - B^z(T - t)) \\ &\quad \times (-\sigma_t^z B^z(S - t)dt + dW_t^z)) \\ &= \sigma^z(t, F_t^z(T, S)) (-\zeta_t^{z,S} dt + dW_t^z) \\ &= \sigma^z(t, F_t^z(T, S)) dW_t^{z,S},\end{aligned}$$

where

$$\sigma^z(t, v) := \sigma_t^z \left[\frac{1}{\Delta} + v \right] ((B^z(S - t) - B^z(T - t))). \quad (30)$$

Using equation (11) and the independence between P^x, P^y ,

$$\frac{P_t^r(S)}{M_t^r} = \frac{P_t^x(S)}{M_t^x} \frac{P_t^y(S)}{M_t^y} = P_0^r(S) \mathcal{E} \left(\int_0^t \sigma_s^x dW_s^x + \int_0^t \sigma_s^y dW_s^y \right),$$

where $\mathcal{E}$ is the Doléans–Dade exponential. Hence,

$$\mathbb{E} \left[\frac{dQ^S}{dQ} \middle| \mathcal{F}_t \right] = \frac{P_t^r(S)}{P_0^r(S)M_t^r} = \mathcal{E} \left(\int_0^t \zeta_s^{x,S} dW_s^x + \int_0^t \zeta_s^{y,S} dW_s^y \right).$$

Applying the Girsanov theorem, the processes $W^{x,S}$ and $W^{y,S}$, defined as $W_t^{z,S} := W_t^z - \int_0^t \zeta_s^{z,S} ds$, $z \in \{x, y\}$, are independent $\mathbb{Q}^S$ -Brownian motions. Hence, $F^z(T, S)$ is a $\mathbb{Q}^S$ -martingale for $z \in \{x, y\}$. Since $\sigma_t^y = \sigma_t^{y^\phi}$ and $B^y = B^{y^\phi}$, $F^{y^\phi}(T, S)$ is also driftless under $\mathbb{Q}^S$. Finally, the martingale property of $F^r(T, S)$ is obtained by applying Ito's product rule to equation (13) together with the independence between $W^{x,S}$ and $W^{y,S}$:

$$\begin{aligned}dF_t^r(T, S) &= (1 + \Delta F_t^y(T, S)) dF_t^x(T, S) \\ &\quad + (1 + \Delta F_t^x(T, S)) dF_t^y(T, S) \\ &\quad + \Delta \underbrace{d\langle F^x(T, S), F^y(T, S) \rangle_t}_{=0}.\end{aligned}$$

This shows that $F^r(T, S)$ is a sum of two $\mathbb{Q}^S$ -martingales and thus itself is a martingale under $\mathbb{Q}^S$, concluding the proof of point (b).

The expression of the diffusion coefficients for x, y and y^ϕ is provided in equation (30). For the Vasicek model, substituting $\sigma_t^x = \sigma_1$ leads to the expression for $\sigma^x(t, F_t^y(T, S))$. For the CIR model, $\sigma_t^y = \sigma_2 \sqrt{y_t}$, which can be written as a function of $F_t^y(T, S)$ using equations (4) and (12). Combining these elements leads to the expression for $\sigma^y(t, F_t^y(T, S))$. Finally, $\sigma^{y^\phi}(t, F_t^{y^\phi}(T, S))$ can be derived using $\sigma_t^{y^\phi} = \sigma_t^y$, $B^{y^\phi} = B^y$ and $A^{y^\phi}(\tau) = e^{-\phi\tau} A^y(\tau)$, concluding the proof of point (c).