ASYMMETRIC SHORT-RATE MODEL WITHOUT LOWER BOUND

Frédéric Vrins, Linqi Wang

LIDAM Discussion Paper LFIN
2021 / 06

UCLouvain

LiDAM
LFIN | Louvain Institute of Data Analysis and Modeling in economics and statistics

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[a] and Linqi Wang[a]

[a]UC Louvain, Louvain Institute of Data Analysis and Modelling in economics and statistics (LIDAM), Louvain Finance Center (LFIN). Voie du Roman Pays 34, B-1348 Louvain-la-Neuve, Belgium.

Abstract

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 Kullback-Leibler (KL) 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 the KL 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: Finance, affine short-rate model, negative interest rates, Kullback-Leibler divergence, implied density calibration

JEL classification: **C52, C61, E43, G12, G13**

*The authors are grateful to Damiano Brigo and Donatien Hainaut for insightful comments on earlier versions of this manuscript. Funding: This work was supported by the Belgian Federal Science Policy Office [ARC grant number 18-23/089]. Emails: frederic.vrins@uclouvain.be (Frédéric Vrins); linqi.wang@uclouvain.be (Linqi Wang).

Highlights

- After two decades of downward trend, interest rate entered the negative territory.
- Standard benchmark models fall short in capturing the salient features of the negative interest rate environment, e.g. asymmetric distribution of interest rates or absence of a strict lower bound, within a unified framework.
- We propose a simple and tractable asymmetric short-rate model that does not display a strict lower bound.
- A new calibration method is introduced which consists in matching market and model implied densities using Kullback-Leibler divergence.
- We analyze the performance of our model in fitting caplet prices and in modelling the forward rate density based on real market data.
- Empirical results show that our model and calibration procedure display superior performances compared to standard benchmarks.

1 Introduction

The Global Financial Crisis of 2007-2008 prompted major central banks to cut their policy rates in order 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. A notable exception is the linear-rational term structure framework introduced in Filipović et al. (2017) for the pricing of swaps and European swaptions and extended in Filipović and Kitapbayev (2018) to American swaption pricing.[1]

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 favored 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 in order to better reflect the recent interest rate environment. The most direct solution would be returning to models exhibiting Gaussian 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

[1] 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.

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 Kullback-Leibler (KL) 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 absolute error (MAE hereafter) minimization is that our approach analyzes 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.

Third, we study the evolution of the market-implied EURIBOR 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 derivatives 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 US. 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 the Hull-White, the CIR++ and the VaCIR++ models during the recent period of negative interest rates in the euro area. We then evaluate the performance of these models when calibrated using both the volatility MAE 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 KL divergence calibration approach provides a more robust way of calibrating interest rate models to fixed-income derivatives 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 the 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 $\phi : \mathbb{R}^+ \to \mathbb{R}$, one obtains a short-rate model that is analytically tractable and able to perfectly fit to 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 benchmark, namely the Vasicek model (noted x hereafter) and the CIR model (noted y hereafter). The short-rate dynamics are respectively given by

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

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

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[2]

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

where W is a Brownian motion. We denote the parameters as $\Psi = (a, b, c, d, z_0)$. 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 $\Psi^y := (\kappa_2\theta_2, -\kappa_2, 0, \sigma_2^2, y_0)$.[3]

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,

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

where A^z, B^z are deterministic functions satisfying Riccati equations, for which analytical solutions can be found.[4]

When $\mathbb{Q}$ stands for the risk-neutral measure, $\mathbb{F}$ is the information 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 $T \geq t$:

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

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.

The existence of analytical expressions for zero-coupon bond prices explains why the Vasicek and CIR models are so popular for interest rates 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

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

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

[3]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).

[4]They are recalled in Appendix A for $z \in \{x, y\}$.

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

$$\phi^z(t) = f_0^M(t) - f_0^z(t) \; , \tag{7}$$

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

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

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:

$$r_t := x_t + y_t \; , \tag{10}$$

where the Brownian motions W^x, W^y in Equation (1) and Equation (2) are assumed to be independent. This assumption preserves the analytical tractability of zero-coupon bond prices[5]

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

[5]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 C for details. We thus restrict ourselves to the independence assumption in this paper.

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).

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[6]

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

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

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

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

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

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

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

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

$$dF_t^x(T,S) = \sigma^x\left(t, F_t^x(T,S)\right) dW_t^{x,S} \ , \ dF_t^y(T,S) = \sigma^y\left(t, F_t^y(T,S)\right) dW_t^{y,S} \ , \tag{14}$$

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

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

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

Proof: see Appendix D.

Because these processes are independent, the conditional density of r_t given (x_s, y_s) is obtained by convolution of the corresponding conditional densities for the Vasicek and the 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 ,*

$$f^S_{z_T^\phi | z_t^\phi}(u) = f^S_{z_T|z_t}(u - \phi^z(t)) \,. \tag{15}$$

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

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

where φ is the standard Normal density.

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

$$f^S_{y_T|y_t}(u) = f_{\chi^2(\nu,\delta(t,T,S))/q(t,T,S)}(u) = q(t,T,S) f_{\chi^2(\nu,\delta(t,T,S))}(q(t,T,S)u) \,, \quad u \geq 0 \,, \tag{17}$$

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

$$q(t,T,S) := 2\left(\rho(T-t) + \frac{\kappa_2 + h}{\sigma_2^2} + B^y(\Delta)\right) ,$$
$$\delta(t,T,S) := \frac{4\rho(T-t)^2 y_t e^{h(T-t)}}{q(t,T,S)} ,$$
$$\rho(\tau) := \frac{2h}{\sigma_2^2(\exp[h\tau]-1)} ,$$

with $\nu := 4\kappa_2\theta_2/\sigma_2^2$ 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

$$f^S_{r_T|(x_t,y_t)}(u) = \int_0^\infty f^S_{x_T|x_t}(u-v) f^S_{y_T|y_t}(v)dv \,. \tag{18}$$

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

$$f^S_{P_T^{z^\phi}(S)|z_t^\phi}(u) = e^{\int_t^T \phi^z(s)ds} f^S_{P_T^z(S)|z_t}\left(u\; e^{\int_t^T \phi^z(s)ds}\right), \quad u \ge 0 \,. \tag{19}$$

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 ($z = y$) are given by

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

where $u \ge 0$ for the Vasicek model and $0 \le u \le 1$ for the CIR model.

As for the combined model, it reads as

$$f^S_{P_T^r(S)|\mathcal{F}_t}(u) = \int_0^1 \frac{1}{v} f^S_{P_T^x(S)|x_t}(u/v) f^S_{P_T^y(S)|y_s}(v)dv, \quad u \ge 0 \,. \tag{21}$$

(b) the density of the spot rate $L^{z^\phi}(T,S) := F_T^{z^\phi}(T,S)$ conditional upon $\mathcal{F}_t = \mathcal{F}_t^x \vee \mathcal{F}_t^y$, [7] *$z \in \{x,y,r\}$, is given by*

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

[7] 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 $L^r(T,S)|(x_t,y_t)$.

3 Model calibration

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

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

where $P_t(S)$ is the price of the zero-coupon bond with maturity S at time t and $f^S_{L(T,S)}$ 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 $f^S_{L(T,S)}$ by $f^S_{L^z(T,S)}$ in Proposition (3) for each of the considered models. Calibration based on the minimization of the mean absolute error of implied volatilities is often favored 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, $f^S_{L^z(T,S)}$, and the market-implied one, $f^S_{L(T,S)}$.

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 absolute error of implied volatilities. We thus calibrate each of the considered models by looking for the parameters Ψ^z which minimize the MAE between the model-implied and market-implied volatilities:

$$\Psi^{z^*} := \arg\min_{\Psi^z}\frac{1}{n}\sum_{i=1}^{n}\left|\sigma_i^M - \sigma_i^z\right| \tag{24}$$

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.[8] Indeed, differentiating twice the expression for the caplet price in Equation (23) with respect to the strike K, we obtain

$$f^S_{L(T,S)|\mathcal{F}_t}(K) = \frac{\frac{\partial^2}{\partial K^2}\mathrm{Cpl}_t(T,S,K)}{\Delta P_t(S)} \ . \tag{25}$$

In order to get an accurate estimation of the density based on Equation (25), we need to obtain the market discount curve 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 in order to obtain a valid (i.e., non-negative) and smooth density.

To this end, we use the arbitrage-free smoothing technique of implied volatility surface developed by Fengler (2009) to interpolate the caplet prices.[9] The procedure takes the quoted caps data and obtain the caplet prices via caplet stripping. Then, a cubic spline smoothing technique is applied to caplet prices, together with no-arbitrage constraints. The interpolated caplet prices are obtained by solving a transformed quadratic program. By eliminating arbitrage opportunities, this approach ensures that the forward rate density is well specified. A detailed outline of this procedure is provided in Appendix B.

In order 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. This problem has been extensively studied in the field of information theory and signal processing, see Basseville (2013) for a review. We thus select as criterion the Kullback-Leibler (KL) divergence.[10] Specifically, the KL divergence between two densities (p, q) is the relative entropy, and is based on (differential) Shannon's entropy:

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

where p is assumed to be absolutely continuous with respect to q.[11]

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

[9] This smoothing technique imposes strict no butterfly-arbitrage restrictions.

[10] The KL divergence is not a distance because the triangular inequality may fail to hold.

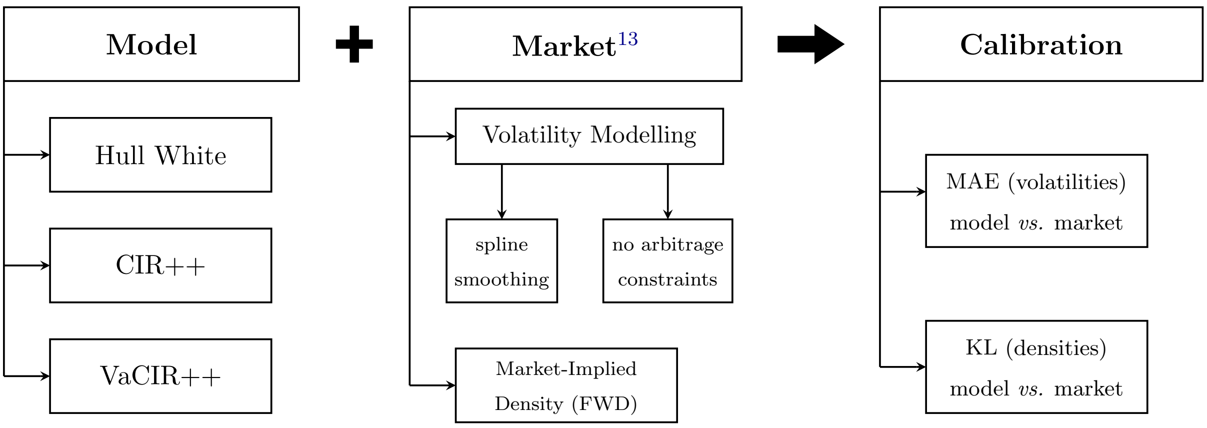
[11] If p is not absolutely continuous with respect to q, we simply restrict the integration domain in Equation (26) to the subset of $\mathbb{R}$ where q is strictly positive.

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

$$\Psi^{z^*} := \arg\min_{\Psi^z} \quad \langle f^S_{L(T,S)} | f^S_{L^z(T,S)} \rangle + \langle f^S_{L^z(T,S)} | f^S_{L(T,S)} \rangle \tag{27}$$

where $f^S_{L(T,S)}$ and $f^S_{L^z(T,S)}$ are respectively the market- and model-implied forward rate densities. Note that we consider a symmetric version of the KL divergence as the relative entropy of the market-implied forward rate density with respect to the model-implied one, the first term on the RHS of Equation (27), will usually be different from the relative entropy of the model-implied forward rate density with respect to the market-implied one, the second term on the RHS of Equation (27). As a consequence, basing the optimization program on only one of them would result in different outcomes.[12]

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 $K \in \{K_0, K_1, K_2 \cdots K_{N-1}, K_N\}$ for certain lower and upper bounds of K_0 and K_N.

---

[12]The interested reader is referred to Chapter 3 of Goodfellow et al. (2016) for more details on the asymmetry of the KL divergence.

[13]Volatility modelling and market-implied density are only needed under the calibration method using KL divergence. The calibration method minimizing the mean absolute 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 absolute 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 B.
- Obtain the market-implied density over a fine grid according to Equation (25).

3: Use KL divergence as the target criterion to calibrate the three models by matching the market-implied density according to Equation (27).

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 KL divergence where the modelled forward density is matched to the market-implied one.

4.1 Data description and volatility curve smoothing

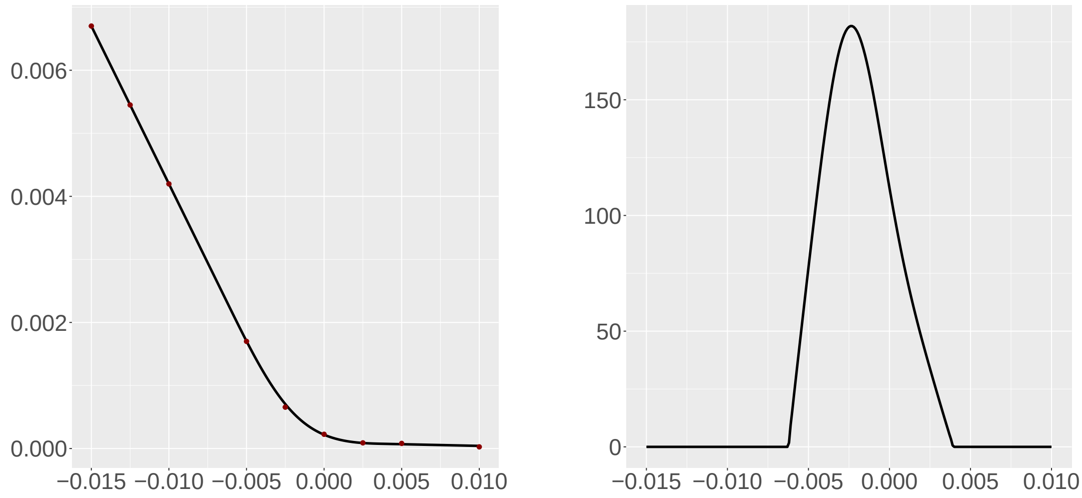
We use historical data for the 6-month EURIBOR forward rate $F_0(6M, 1Y)$ and the interest rate derivatives whose underlying rate is the 6-month forward rate in 6 months, i.e. $L(6M, 1Y) = F_{6M}(6M, 1Y)$.[15] Observe that the market provides quotations for the implied volatilities of caps whereas, in order 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 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.

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

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.



(a) Interpolated price curve.

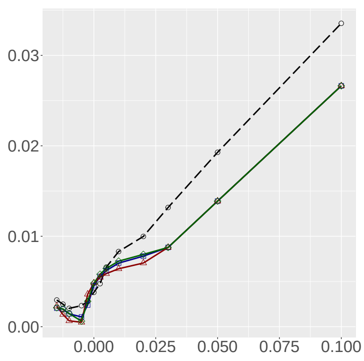
(b) Implied density of $L(6M, 1Y)$.

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 Bachelier formula alongside with interpolated price curve obtained by 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).

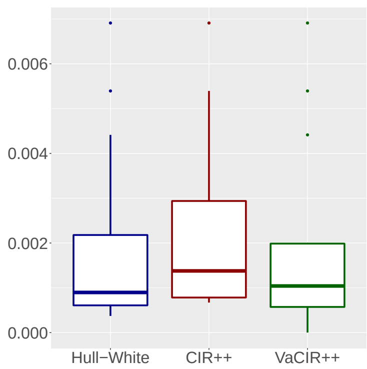
In Figure 1, we illustrate the performance of the smoothing approach from two perspectives, the fit of market prices and the behavior of the probability density. Figure 1a presents the interpolated prices and the market quoted prices for the caplet products of interest, $\text{Cpl}(0, 6M, 1Y)$. It can be observed that all market quotations are in line with the smoothed price curve, indicating that the interpolation technique is performing well as an accurate caplet valuation tool. Figure 1b 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

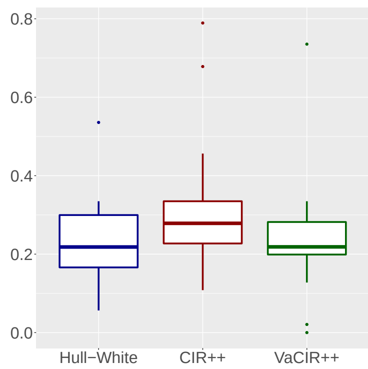
Equation (23) provides the model-implied caplet prices. The corresponding caplet volatilities, obtained via Bachelier formula, can thus be compared with those extracted from market quotes as explained in Section 3.1. The model is then calibrated according to Equation (24). We provide the calibration performance on the first available date from our sample as an illustration.



(a) Implied volatility curve.



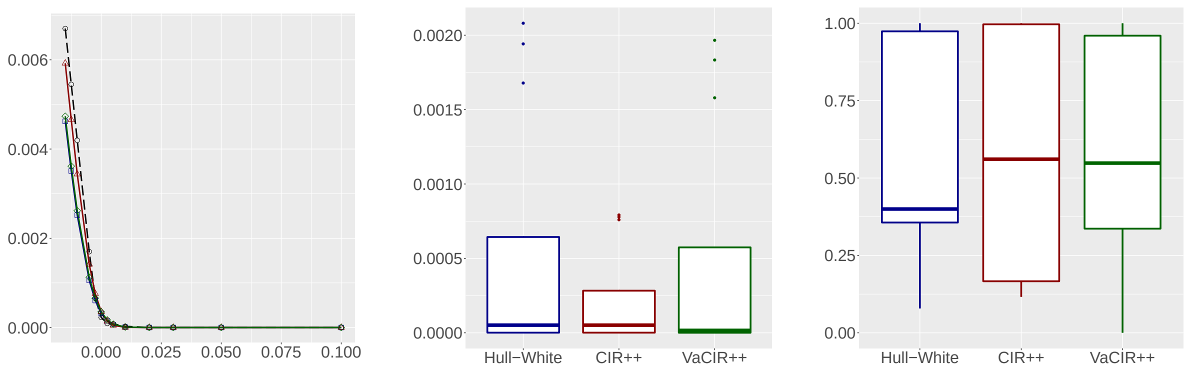
(b) Implied vol absolute error.



(c) Implied vol relative error.

Figure 2: Implied volatility curves with associated absolute and relative volatility errors on 31 March 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 (□), CIR++ (△) and VaCIR++ (⋄) models. Panel (b) and Panel (c) provide respectively the box plots of the associated absolute and relative errors for the three candidate models.

We plot in Figure 2 the modelled volatilities with associated absolute and relative errors using the parameters calibrated based on fitting the market-implied volatilities. Figure 2a 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 all three considered 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 box plots of the associated absolute and relative errors, displayed in Figure 2b and Figure 2c respectively, provide further details on the relative performance of the three considered candidates. The CIR++ model has the largest median in terms of the volatility absolute and relative errors while the error distribution of the VaCIR++ model is close to the Hull-White model but exhibits less dispersion. On this date, the Hull-White model outperforms the CIR++ model, but this is not always the case. For instance, the opposite situation happens on 29 January 2021 (see Figure 10 in Appendix E). The persistent outperformance of the VaCIR++ model over both component processes illustrates the ability of the framework to accommodate different circumstances and to deliver a better fit of the market data.



(a) Caplet price curve.

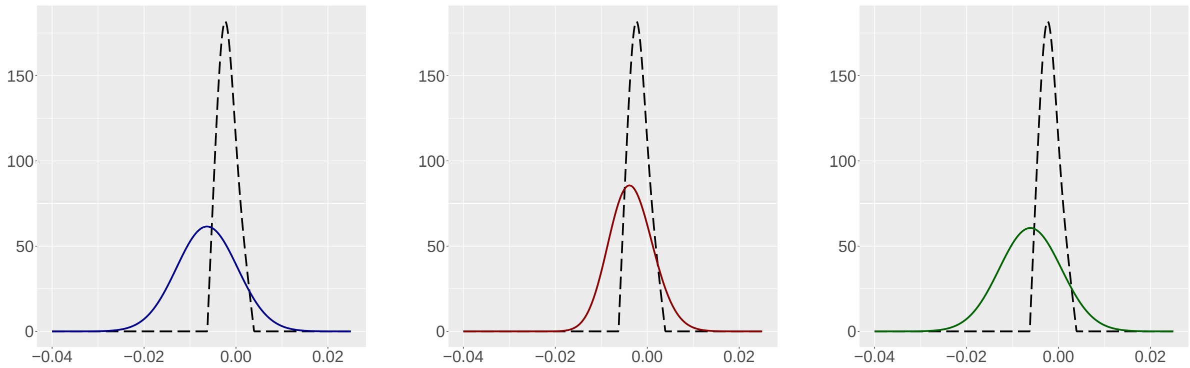
(b) Price absolute error.

(c) Price relative error.

Figure 3: Caplet price curves with corresponding absolute and relative pricing errors on 31 March 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 (□), CIR++ (△) and VaCIR++ (⋄) models. Panel (b) and panel (c) provide respectively the box plots of the corresponding absolute errors and relative errors for the three candidate models.

Figure 3 shows the same results as Figure 2 but expressed in terms of prices rather than volatilities. It can be observed that although the Hull-White model provides a better fit for the implied volatility curve, this does not directly translate into higher pricing accuracy. In particular, the Hull-White model has the highest value in both absolute and relative errors while the CIR++ model and the VaCIR++ model provide, respectively, the best performance in terms of the absolute and the relative errors.

Not surprisingly, Figure 3a 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.



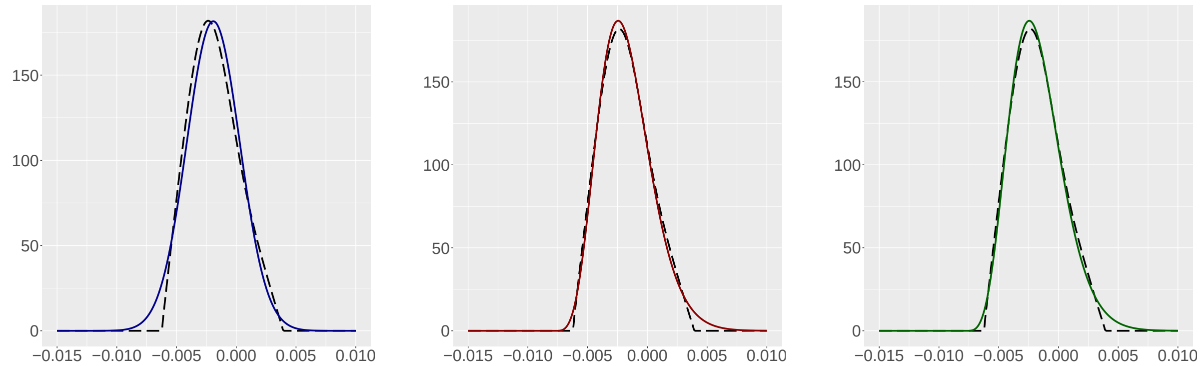
(a) Hull-White model. (b) CIR++ model. (c) VaCIR++ model.

Figure 4: Density curves of the forward rate $L(6M, 1Y)$ on 31 March 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++ model, respectively (solid lines).

We further investigate the distribution of the forward rate under the three considered models compared to the market-implied one using the same parameters. Figure 4 displays respectively the density curves modelled by the Hull-White, the CIR++ and the VaCIR++ models against the market-implied one. We observe that in general, the three models have a higher peak and their densities approach zero at a slower speed when the strikes decrease to deeply negative values or increase to highly positive rates. Moreover, it can be observed from Figure 4a to Figure 4c that the densities under all three models, compared to the market density, are more flat with a shift to the left side. This observation suggests that the calibration approach focusing on a few data points of the implied volatilities does not guarantee the match of the entire forward rate distribution.

4.3 Calibration results for the forward rate density

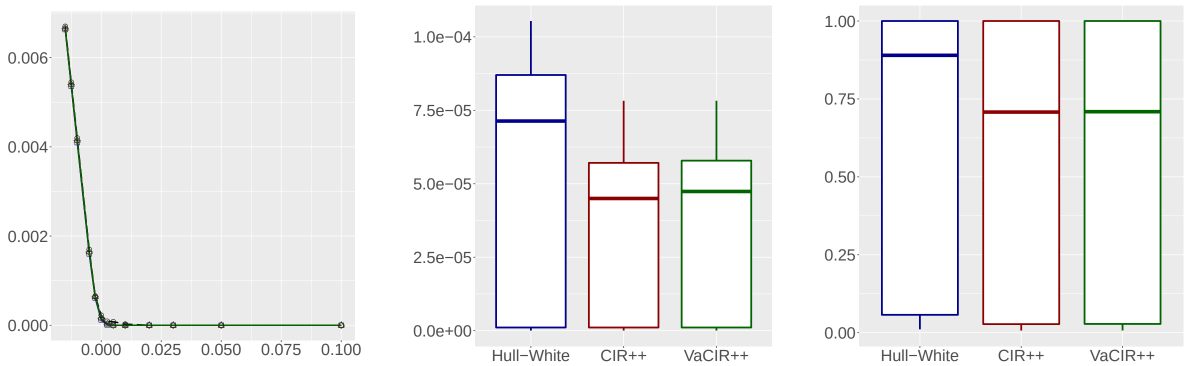
Equipped with the smoothed caplet price curve computed over a fine grid of strikes, we now turn to the estimation of the market-implied forward rate distribution according to Equation (25). We then calibrate the three considered models according to Algorithm 2 to perfectly match the market-implied density by minimizing the distance (measured by KL divergence) between the model-implied and the market-implied curves. 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.



(a) Hull-White model. (b) CIR++ model. (c) VaCIR++ model.

Figure 5: Density curves of the forward rate $L(6M, 1Y)$ on 31 March 2016 where models are calibrated on the criteria of KL 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, the CIR++ and the VaCIR++ model, respectively (solid lines).

We display in Figure 5 the density curves modelled by the three frameworks (solid lines). We observe that all three densities approach zero at a slower speed than the market-implied density when the strikes move toward extreme negative or positive values. More specifically, Figure 5a shows that the modelled 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, the densities under the CIR++ and the VaCIR++ models, exhibited in Figure 5b and Figure 5c, have a slightly higher peak. Both models match the market density remarkably well except in the right tail and they capture the overall market distribution curve better than the Hull-White model. Furthermore, the VaCIR++ model substantially outperform the Hull-White model in terms of KL divergence, but only marginally so for the CIR++ process. This observation is consistent with the argument that the VaCIR++ model will perform at least as well as its 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 MAE criterion.



(a) Caplet price curve.

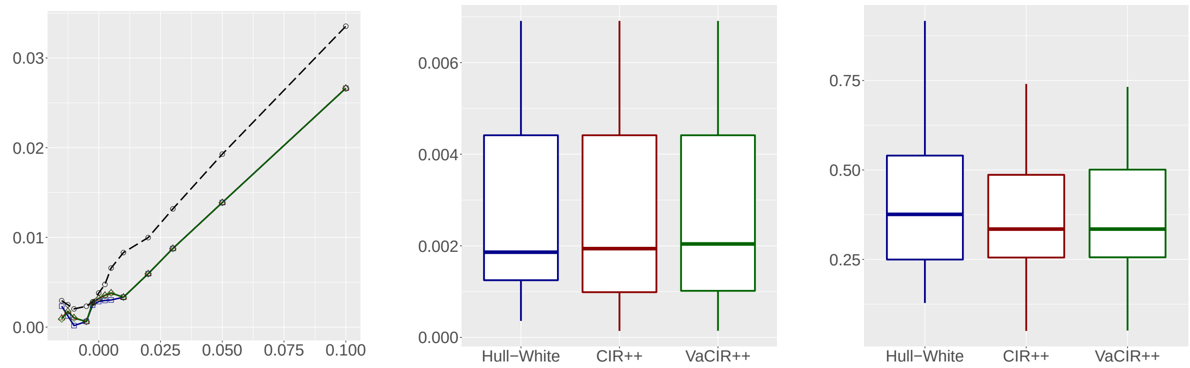
(b) Price absolute error.

(c) Price relative error.

Figure 6: Caplet price curves with corresponding absolute and relative pricing errors on 31 March 2016 where models are calibrated on the criteria of KL divergence minimization. Panel (a) displays the market and modelled caplet prices with respect to the strike for the market (dashed), Hull-White (□), CIR++ (△) and VaCIR++ (⋄) models. Panel (b) and panel (c) provide respectively the box plots of the corresponding absolute errors and relative errors for the three candidate models.

We further illustrate in Figure 6 the pricing accuracy of the three candidates using the model parameters calibrated by density matching. Figure 6a displays the modelled caplet prices using the three considered candidate models compared to the market data. It can be concluded that all three models are able to provide a quite accurate valuation for the caplet products. We plot the absolute and relative pricing error in Figure 6b and Figure 6c, and we can draw a similar conclusion about relative model performance as in the case of density comparison. More specifically, the Hull-White model exhibits the lowest pricing accuracy among the three candidates, the performance of the CIR++ and the VaCIR++ models are superior and rather close to each other since the CIR process receives a larger weight in the VaCIR++ model.

Comparing panels (b) and (c) of Figure 3 and Figure 6, we observe that the better density fit obtained when adopting the KL divergence calibration technique instead of the implied volatility MAE criterion does not come at the cost of larger pricing errors. On the contrary, the dispersion of the pricing errors is substantially reduced when using the density matching calibration based on the KL divergence.



(a) Implied volatility curve. (b) Implied vol absolute error. (c) Implied vol relative error.

Figure 7: Implied volatility curves with associated absolute and relative volatility errors on 31 March 2016 where the models are calibrated on the criteria of KL divergence minimization. Panel (a) contains the implied volatility curves with respect to the strike for the market (dashed), Hull-White (□), CIR++ (△) and VaCIR++ (◇) models. Panel (b) and Panel (c) provide respectively the box plots of the associated absolute errors and relative errors for the three candidate models.

Additionally, we demonstrate in Figure 7 the model performance in terms of how close the model-based implied volatility curves, arising from minimizing the KL divergence, are to the market-based implied volatility curve. Figure 7a shows that the CIR++ and VaCIR++ models have an overall better performance than the Hull-White process for both in and at the money caplet products while the three frameworks are indistinguishable when evaluated at the deep out of the money range.

Furthermore, if we compare these three plots with the corresponding plots in Figure 2, we observe that calibration focused on density matching has slightly higher but comparable implied volatility absolute and relative errors compared to the calibration results obtained by directly fitting the implied volatility curve. 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.

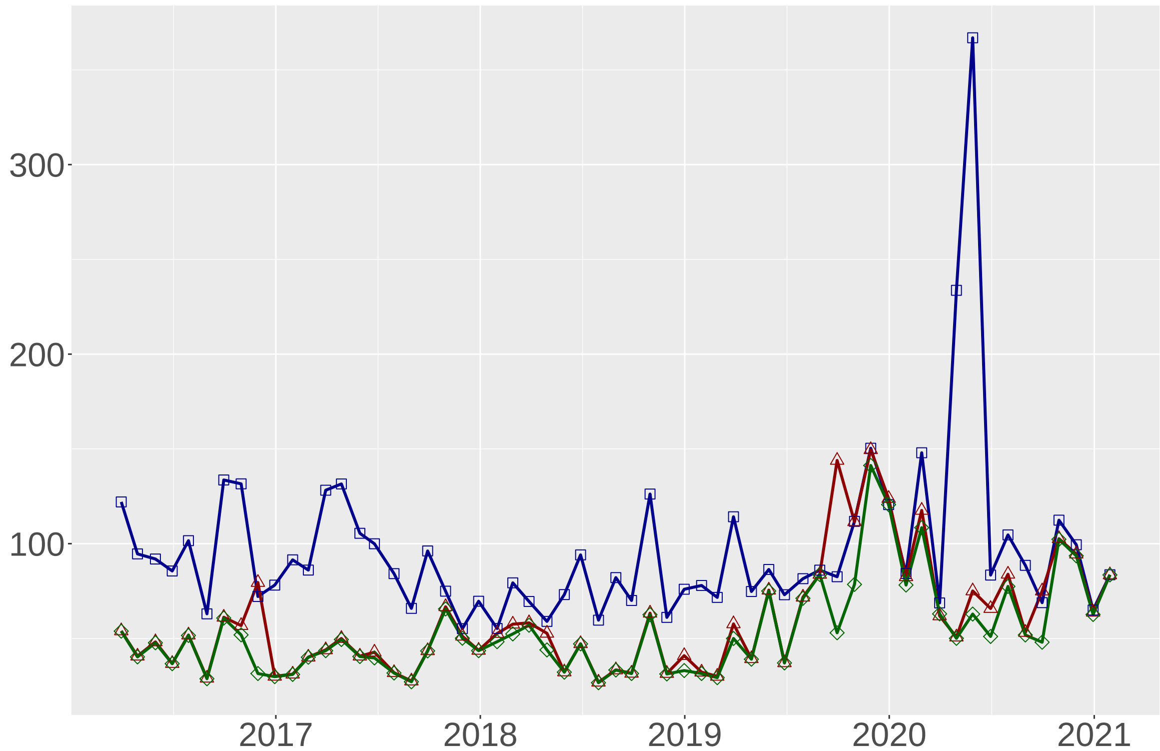
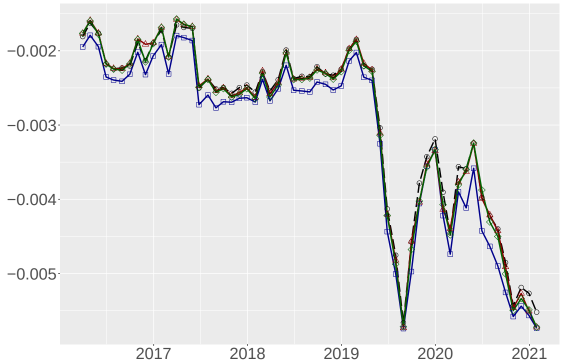
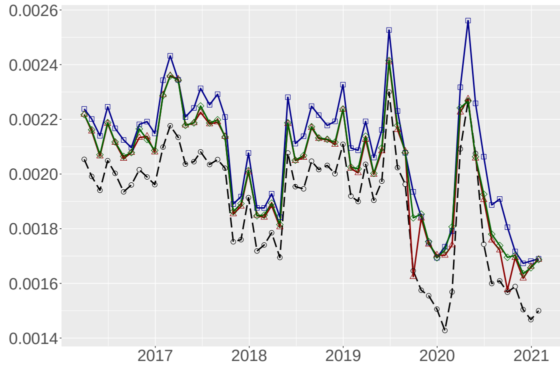


Figure 8: Evolution over time of the KL divergence resulting from the density matching calibration under the Hull-White (□), CIR++ (△) and VaCIR++ (⋄) models.

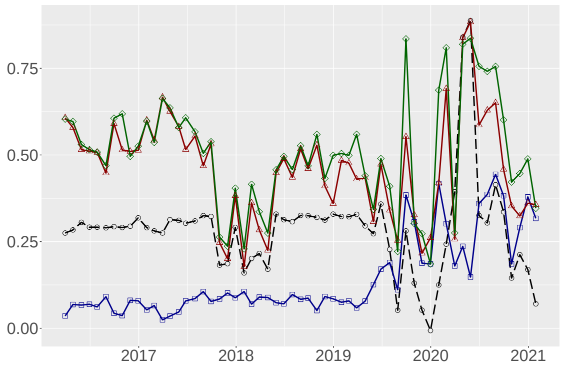
We further evaluate the performance of the considered models in terms of matching the market-implied density curve over time. The candidate models are calibrated on a monthly basis and we report in Figure 8 the KL divergence under the three frameworks. The best-in-class property of VaCIR++ is preserved through time. In particular, it always has the closest matching to the market density curve while the worst performing model is alternating between the Hull-White model and the CIR++ model. Also, we can observe that the CIR++ model, in general, has a better performance than the Hull-White model which makes the CIR++ model an attractive candidate as a single-factor model with the shift providing the ability to deal with negative rates. To summarize, the VaCIR++ model is more advantageous both theoretically - allowing for negative values, skewness, and no strict lower bound - and empirically as it enhances the fitting performance and provides more stability in the calibration over time compared to both components.



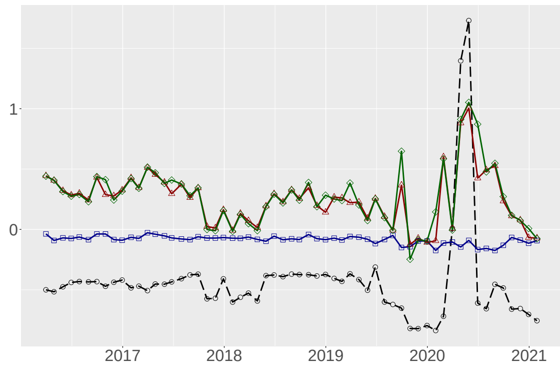
(a) Conditional mean.



(b) Conditional standard deviation.



(c) Conditional skewness.



(d) Conditional excess kurtosis.

Figure 9: Time series of the $L(6M, 1Y)$ conditional moments for the market ($\circ$) and for the Hull-White ($\square$), CIR++ ($\triangle$) and VaCIR++ ($\diamond$) models.

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.[16] In this manner, we compute the first four conditional moments for the 6-month forward rate in 6 months for the whole sample period. We provide further investigation on the detailed performance of the three competing models compared to the market-implied moments. The results are presented in Figure 9. It can be observed from Figure 9a that the CIR++ and the VaCIR++ models are able to track the market-implied mean accurately while the mean under the Hull-White model diverges from the market-implied one. Figure 9b shows that the Hull-White has the

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

worst performance in terms of tracking the market standard deviation while the closet fitted framework is alternating between the CIR++ model and the VaCIR++ model. For the skewness illustrated in Figure 9c, the CIR++ and VaCIR++ models provide more skewness than the market-implied one, but overall a better fit than the Hull-White model which always exhibits insufficient skewness in the distribution. In addition, Figure 9d illustrates that the Hull-White model deviates less from the market kurtosis but it fails to adjust for extreme market conditions after 2020, compared to the CIR++ and the VaCIR++ frameworks.

In addition, focusing on the forward rate density and moments provides us with more insights from the market which can be crucial for risk management and hedging applications. As shown in Figure 9a and Figure 9b, 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 observation indicates that during the COVID-19 period, the interest rate displays a more volatile progression. As a consequence, higher volatility is generated due to the increased uncertainty in financial markets and concerns for funding conditions due to the crisis. Furthermore, Figure 9c shows that the conditional skewness is generally positive with a decrease during the COVID-19 crisis period. We also observe from Figure 9d a negative excess kurtosis indicating a lower tail-thickness compared to a Gaussian distribution.

5 Conclusions

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 absolute error. Precisely, the model is calibrated such that the model-implied density is the closest from the market-implied one in terms of Kullback-Leibler 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 absolute error minimization and our proposed density matching approach based on the minimization of the Kullback-Leibler divergence. We notice the outperformance of our model relative to the Hull-White and the CIR++ models. Moreover, the calibration procedure based on the Kullback-Leibler divergence significantly enhances the matching of the forward densities while preserving high pricing accuracy and comparable fitting quality for implied volatility curves.

Our 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.

References

Acharya, V. V., T. Eisert, C. Eufinger, and C. Hirsch (2019). Whatever it takes: The real effects of unconventional monetary policy. *The Review of Financial Studies*, 32.9, 3366–3411.

Basseville, M. (2013). Divergence measures for statistical data processing — An annotated bibliography. *Signal Processing*, 93.4, 621–633.

Bauer, M. D. and M. Chernov (2021). Interest rate skewness and biased beliefs. *NBER Working Paper*, No. 28954.

Breeden, D. T. and R. H. Litzenberger (1978). Prices of state-contingent claims implicit in option prices. *The Journal of Business*, 621–651.

Brigo, D. and F. Mercurio (2001). A deterministic–shift extension of analytically–tractable and time–homogeneous short–rate models. *Finance and Stochastics*, 5.3, 369–387.

Brigo, D. and F. Mercurio (2007). Interest rate models-theory and practice: With smile, inflation and credit. Springer Science & Business Media.

Cox, J. C., J. E. Ingersoll Jr, and S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica*, 53.2, 385–407.

Fengler, M. R. (2009). Arbitrage-free smoothing of the implied volatility surface. *Quantitative Finance*, 9.4, 417–428.

Filipović, D., M. Larsson, and A. B. Trolle (2017). Linear-rational term structure models. *The Journal of Finance*, 72.2, 655–704.

Filipović, D. and A. B. Trolle (2013). The term structure of interbank risk. *Journal of Financial Economics*, 109.3, 707–733.

Filipović, D. and Y. Kitapbayev (2018). On the American swaption in the linear-rational framework. *Quantitative Finance*, 18.11, 1865–1876.

Goodfellow, I., Y. Bengio, and A. Courville (2016). Deep Learning. MIT Press.

Hagan, P. and M. Konikov (2004). Interest rate volatility cube: Construction and use. *Bloomberg technical report*, No. 62.

Heider, F., F. Saidi, and G. Schepens (2019). Life below zero: Bank lending under negative policy rates. *The Review of Financial Studies*, 32.10, 3728–3761.

Hull, J. and A. White (1990). Pricing interest-rate-derivative securities. *The Review of Financial Studies*, 3.4, 392–573.

Li, H. and F. Zhao (2009). Nonparametric estimation of state-price densities implicit in interest rate cap prices. *The Review of Financial Studies*, 22.11, 4335–4376.

Longstaff, F. A., S. Mithal, and E. Neis (2005). Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market. *The Journal of Finance*, 60.5, 2213–2253.

Trolle, A. B. and E. S. Schwartz (2014). The swaption cube. *The Review of Financial Studies*, 27.8, 2307–2353.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5.2, 177–188.

A 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

$$
\begin{aligned}
A^x(\tau) &= \exp\left\{\left(\theta - \frac{\sigma^2}{2k^2}\right)[B(\tau) - \tau] - \frac{\sigma^2}{4k} B^x(\tau)^2\right\} , \\
B^x(\tau) &= \frac{1}{k}\left[1 - e^{-k\tau}\right] .
\end{aligned}
$$

For the CIR model specified in Equation (2),

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

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

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

B Arbitrage-free smoothing of the volatility curve

Following the notations in Fengler (2009), assume that we observe the caplet prices $\mathbf{y} = (y_1, \ldots, y_n)$ at the strikes $a = u_0, \ldots, u_{n+1} = b$, 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 $i = 1, \ldots, n$. Furthermore, we define $\mathbf{g} = (g_1, \ldots, g_n)^\top$ and $\boldsymbol{\gamma} = (\gamma_2, \ldots, \gamma_{n-1})^\top$. 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 $i = 1, \ldots, n-1$, and the elements $q_{i,j}$, of matrix $\mathbf{Q}$, for $i = 1, \ldots, n$ and $j = 2, \ldots, n-1$, are given by

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

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

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

$$\begin{aligned} r_{i,i} &= \tfrac{1}{3}(h_{i-1} + h_i) \ , && \text{for } i = 2, \ldots, n-1 \\ r_{i,i+1} &= r_{i+1,i} = \tfrac{1}{6} h_i \ , && \text{for } i = 2, \ldots, n-2 \\ r_{i,j} &= 0 \ , && \text{for } |i - j| \geq 2 \end{aligned}$$

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

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

where $\mathbf{W}_n = \operatorname{diag}(w_1, \ldots, w_n)$. 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

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

C Correlation between two factors

C.1 Introduce correlation winthin the affine framework

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

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

which is equivalent to

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

We can re-write 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, 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

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

The zero coupon bond price is defined as

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

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

$$\begin{aligned}
\frac{dB^1(t,T)}{dt} &= 1+\kappa_2 B^1(t,T) - \frac{1}{2}\left(\sigma_{21}B^2(t,T)+\sigma_{11}B^1(t,T)\right)^2 \ , \\
\frac{dB^2(t,T)}{dt} &= 1+\kappa_1 B^2(t,T) \ , \\
\frac{dA(t,T)}{dt} &= -\kappa_1\theta_1 B^1(t,T) - \kappa_2\theta_2 B^2(t,T) - \frac{1}{2}\sigma_{22}^2 B^2(t,T)^2 \ .
\end{aligned}$$

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.

C.2 Introduce correlation by correlated Brownian Motions

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

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

We can re-write 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 $corr(dW_t^1, dW_t^3) = \rho dt$.

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

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

D 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 $z \in \{x, y\}$

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

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

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

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

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

where

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

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

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

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

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

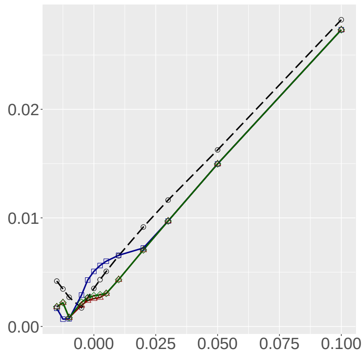
Applying the Girsanov theorem, the processes $W^{x,S}$ and $W^{y,S}$, defined as $W_t^{z,S} := W_t^z - \int_0^t \zeta_s^{z,S}ds$, $z \in \{x, y\}$, are independent $\mathbb{Q}^S$-Brownian motions. Hence, $F^z(T,S)$ is a $\mathbb{Q}^S$-martingale for $z \in \{x, y\}$. Since $\sigma_t^y = \sigma_t^{y^\phi}$ 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 $W^{y,S}$:

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

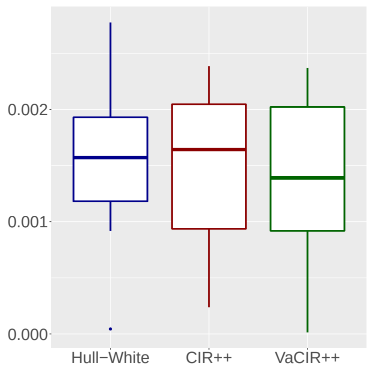
This shows that $F^r(T,S)$ is a sum of two $\mathbb{Q}^S$-martingales and thus itself is a martingale under $\mathbb{Q}^S$, concluding the proof of Point (b).

The expression of the diffusion coefficients for x, y and y^ϕ are provided in Equation (28). 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 Equation (4) and Equation (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).

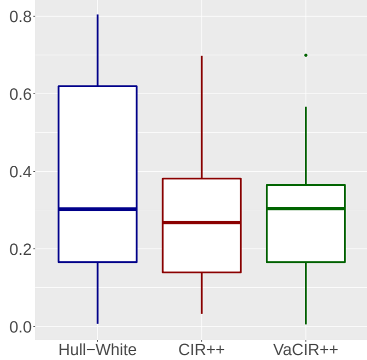
E Calibration results of implied volatility curve: 29 January 2021



(a) Implied volatility curve.



(b) Implied vol absolute error.



(c) Implied vol relative error.

Figure 10: Implied volatility curves with associated absolute and relative volatility errors on 29 January 2021 where the models are calibrated on the criteria of implied volatility error minimization. Panel (a) contains the implied volatility curves with respect to different strike rates. Panel (b) and Panel (c) provide respectively the box plots of the associated absolute errors and relative errors for the three candidate models.

We take the calibration results on 29 January 2021 to illustrate the performance of the three frameworks where the CIR++ model has superior performance compared to the Hull-White model. The implied volatility curves with the corresponding errors are exhibited in Figure 10. We note that although the CIR++ model has a slightly higher median absolute error compared to the Hull-White model, it has an overall smaller discrepancy from the market data, especially evaluated in terms of relative volatility errors. Moreover, the VaCIR++ model further reduces the absolute pricing error and its performance is closer to the CIR++ framework.