Economic Modelling 57 (2016) 180–196

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.elsevier.com/locate/ecmod

A bivariate Hawkes process for interest rate modeling

Donatien Hainaut

ESC Rennes Business School & CREST, France

ARTICLE INFO

Article history:
Received 25 August 2015
Received in revised form 19 April 2016
Accepted 20 April 2016
Available online 12 May 2016

Keywords:
Hawkes process
Self and mutually exciting processes
Interest rate
Micro-structure
Yield curve

ABSTRACT

This paper proposes a continuous time model for interest rates, based on a bivariate mutually exciting point process. The two components of this process represent the global supply and demand for fixed income instruments. In this framework, closed form expressions are obtained for the first moments of the short term rate and for bonds, under an equivalent affine risk neutral measure. European derivatives are priced under a forward measure and a numerical algorithm is proposed to evaluate caplets and floorlets. The model is fitted to the time series of one year swap rates, from 2004 to 2014. From observation of yield curves over the same period, we filter the evolution of risk premiums of supply and demand processes. Finally, we analyze the sensitivity of implied volatilities of caplets to parameters defining thelevel of mutual-excitation.
© 2016 Elsevier B.V. All rights reserved.

1. Introduction

During the recent crisis of European sovereign debts, fixed income markets collapsed and caused liquidity shortfalls in countries of South Europe. The immediacy of information contributed to speed up the tightening of traded volumes of short and long term bonds. And the abrupt decline in demand for debts, due to the anxiety about excessive national debt, even if correlated with a reduction of supply, raised interest rates to historical summits, in Greece (33.7% for the 10 year bond on the 3/2/2012), Italy, Spain and Portugal. On the other side,by the end of 2011, Germany was estimated to have made more than €9 billion out of the crisis as investors flocked to safer but near zero interest rate German federal government bonds. By July 2012 the Netherlands, Austria and Finland also benefited from zero or negative interest rates, as consequence of the high demand for their national debt. This crisis reminds us that interest rates basically depend on the law of supply and demand. There is also compelling evidence that yields of fixed income instruments are affected by liquidity concerns, as shown by Landschoot (2008), Longstaff (2004), Chen et al. (2007), Covitz and Downing (2007) or Acharya and Pedersen (2005). Understanding liquidity effects in bonds markets is then of particular importance for central banks, to define appropriate monetary policy actions.

As liquidity shortages result from a disequilibrium between the global demand and supply for debts, the model developed in this work assumes that the short term rate is ruled by a bivariate Hawkes processes, representing the aggregate bid and ask orders, for fixed income instruments. This approach is fully relevant with the monetary theory as e.g. detailed in the chapter 5 of Mishkin (2007), and presents several interesting features. Firstly, it introduces path dependency and auto-correlation, that are absent from models based on Brownian motions (Brigo and Mercurio, 2007, Cox et al., 1985, Dai and Singleton, 2000, Duffie and Kan, 1996, Hull and White, 1990, Zhang et al., 2015, for a survey), on Lévy processes (Eberlein and Kluge, 2006, Filipović and Tappe, 2008, Hainaut and MacGilchrist, 2010) or on switching processes (Hainaut, 2013; Shen and Siu, 2013). Secondly, it adds mutual excitation and snowball effects, between the supply and demand in interest rate markets. This features is introduced in the dynamics through a bivariate Hawkes process (see Hawkes, 1971a,b; Hawkes and Oakes, 1974). This is a parsimonious self and mutually exciting point process for which the intensity jumps in response and reverts to a target level in the absence of event. As the future of a Hawkes process is influenced by the timing of past events, Errais et al. (2010) use this, combined with a mean reverting drift of the intensity, to generate contagion between defaults in a top down approach to credit risk. Embrechts et al. (2011) apply multivariate Hawkes processes in their analysis of stocks markets. Mutually exciting processes are also used by Aït-Sahalia et al. (2015, 2014), to model two key aspects of asset prices: clustering in time and cross sectional contamination between regions. Bormetti et al. (2015) model systemic price cojumps with a Hawkes factor models. Rambaldi et al. (2015) propose a Hawkes-process approach to explain foreign exchange market activity around macroeconomic news. Zhu (2014) and Hainaut (forthcoming) study affine models with Hawkes jumps. On the other hand, these processes are increas-

E-mail address: donatien.hainaut@esc-rennes.fr.

http://dx.doi.org/10.1016/j.econmod.2016.04.016
0264-9993/© 2016 Elsevier B.V. All rights reserved.

ingly integrated in high frequency finance. Examples include the modeling of the duration between trades (Bauwens and Hautsch, 2009) or the arrival process of buy and sell orders, as in Bacry et al. (2013). Giot (2005), Chavez-Demoulin et al. (2005) or Chavez-Demoulin and McGill (2012) test these processes in a risk management context. Hawkes processes have also been applied to insurance and ruin theory in Dassios and Zhao (2012), Stabile and Torrisi (2010) or Zhu (2013b). A recent survey of the literature on Hawkes processes applied to finance is proposed in Bacry et al. (2015). Hawkes processes also provide an alternative to model based on the theory of extreme values. This approach (see e.g. Embrechts et al., 1997) is by the way applied in a paper of Bali (2007) to explain interest rate volatility.

This research complements the existing literature about interest rate modeling in several directions. It is one of the first to use exclusively a bivariate Hawkes process for the modeling of the term structure of interest rates. Secondly, the model is built on strong economics foundations. Whereas most of existing models, such as developed by Vasicek (1977) and Hull and White (1990) are based on econometrics and are not necessarily reconcilable with the theory of money economics. Another advantage of the proposed bivariate process is that the two factors driving the interest rates can be assimilated to volumes of supply and demand for fixed income assets. As the calibration is easy and robust, these factors can also serve us to monitor the trading activity in interest rate markets, that are mainly "over the counter" and for which we don't have information about exchanged volumes. Thirdly, this work provides all the tools for pricing bonds and for reconciling the dynamics of the short term rate under the real measure, with the term structure of bonds yields, evaluated under the risk neutral measure. We propose a family of changes of measure that preserves the dynamics of the process under real and risk neutral measures. Finally, after an analysis of the dynamics of bond quotes, the moment generating function of bond yields under a forward measure is detailed and a discrete Fourier transform algorithm is proposed to price derivatives.

The paper proceeds as follows. Section 2 introduces the model and derives its main features, like the moments of intensities for the arrival of bid and ask orders. In Section 3, we present equivalent exponential affine measures and study the conditions ensuring that the equivalent measure is risk neutral. After a presentation of the dynamics of the short term rate under this risk neutral measure, a formula for bond pricing is proposed. The Section 4 is about the valuation of derivatives. In Section 5, we fit the model to the time series of one year swap rate. From observation of yield curves over the same period, we filter the evolution of risk premiums of supply and demand processes. Finally, we test the sensitivity of yield curves and smiles of implied volatilities to changes of parameters.

2. Model

The proposed approach to the analysis of interest rate determination looks at supply and demand in the bond market. It finds its foundations in the economics theory as detailed, for example in the chapter 5 of Mishkin (2007) . Each bond price is associated with a particular level of interest rate. Specifically, the negative relationship between bond prices and interest rates means that when a bond price rises its yield falls and vice versa. In economics, the relationship between the quantity supplied and the price is described by the bond demand line. Under the assumption that all other economic variables are held constant, quantities of supplied bonds, noted B_1, increases linearly with bond prices P_1. The supply curve is then described by the following relation:

$$P_1 = L^1 + \theta_1 B_1 \,,$$

where L^1 and θ_1 are respectively the intercept and the elasticity of the supply curve. Under the same assumption, we can derive a demand curve that shows the relationship between the quantity demanded and bond prices. This curve has the usual downward slope found in demand curves, indicating that as the price increases (everything else being equal), the demanded quantity of bonds falls. The equation defining this line is as follows:

$$P_2 = L^2 - \theta_2 B_2 \,,$$

where L^2 and θ_2 are respectively the intercept and the elasticity of the demand curve. The market equilibrium occurs when the amount of demand equals the supply, $B^* = B_d = B_s$, at a price P^* such that,

$$P^* = L^2 - \theta_2 B^* = L^1 + \theta_1 B^* . \tag{1}$$

In economics, a change in market conditions is represented by a parallel shift in the demand or supply curve. Mathematically, this shift corresponds to a modification of the intercept L^2 or L^1. Then, so as to model the dynamics of interest rates, L^1 and L^2 are indexed by the time, t. According to relation (1), the volume of exchanged bonds at any given time is hence equal to

$$B_t^* = \frac{L_t^2 - L_t^1}{\theta_1 + \theta_2}, \tag{2}$$

and the equilibrium bond price at time t, that is inversely proportional to the equilibrium yield is given by

$$P_t^* = \frac{\theta_1}{\theta_1 + \theta_2} L_t^2 - \frac{\theta_2}{\theta_1 + \theta_2} L_t^1 \propto \frac{1}{r_t}. \tag{3}$$

The formation of this equilibrium and the impact of a marginal shock on the demand curve is illustrated in Fig. 1. As the interest rate is inversely proportional to P^*, if the function $\frac{1}{r}$ is approached by a Taylor development of first order, we infer that the equilibrium interest rate, that is denoted by r_t, is proportional to the difference between the intercepts of supply and demand curves, scaled by some positive constants:

$$r_t \propto \frac{\theta_2}{\theta_1 + \theta_2} L_t^1 - \frac{\theta_1}{\theta_1 + \theta_2} L_t^2 . \tag{4}$$

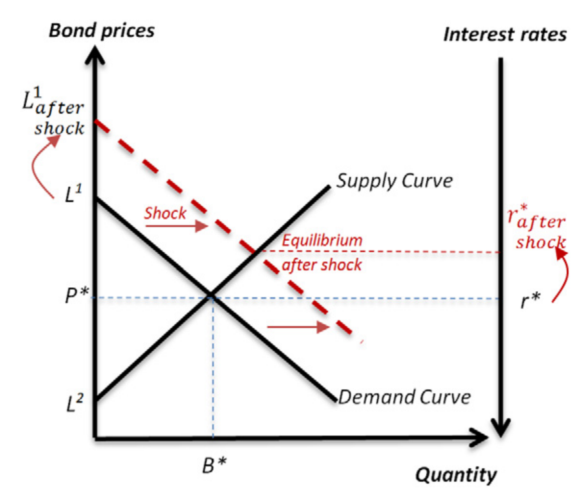

Fig. 1. This graph illustrates the relation between bond prices, interest rates and supply–demand curves. It also shows that a marginal positive shock on the demand corresponds to an increase of the intercept in the demand equation. This positive shock on demand raises bond prices and decreases the equilibrium yield.

According to the monetary theory, the interest rate is the difference between quantities representative of the supply and of the demand. Starting from this theoretical result, we postulate in the remainder of this paper that the short term interest rate, r_t is the sum of a function of time $\varphi(t)$ and of a process X_t that is a difference between supply and demand

$$r_t = \varphi(t) + X_t, \tag{5}$$

with

$$X_t = \alpha_1 L_t^1 - \alpha_2 L_t^2 \tag{6}$$

on a complete probability space $(\Omega, \mathcal{F}, P)$, with a right-continuous information filtration $\mathbb{F} = (\mathcal{F}_t)_{t>0}$, where P denotes from now on the real probability measure. In order to define a realistic micro-structure price model while accounting for the impact of market orders, the framework of multivariate Hawkes processes is well suited. Inspired from the work of Bacry et al. (2013), the supply and demand quantities that rules X_t are related to numbers and sizes of bid-ask orders. These orders and their numbers are respectively noted O^1, O^2 and N_t^1, N_t^2. From now on, L_t^1 and L_t^2 point out the processes modeling the aggregate supply and demand instead of intercepts of demand-supply curves. They are defined as the total of all bid and orders till time t:

$$L_t^1 = \sum_{i=1}^{N_t^1} O_i^1, \tag{7}$$

$$L_t^2 = \sum_{i=1}^{N_t^2} O_i^2. \tag{8}$$

As illustrated in the introduction of this paragraph, an increase of the aggregate offer of bonds causes a decline of their prices and a rise of interest rates. In the opposite scenario, under the pressure of a high aggregate demand, bonds prices grow up and interest rates drop. Then if α_1 and α_2 respectively denotes the permanent impact of sell and buy orders of bonds, the economics theory suggests therefore the following dynamics for X_t:

$$\begin{aligned} dX_t &= \alpha_1 dL_t^1 - \alpha_2 dL_t^2 \\ &= \alpha_1 O^1 dN_t^1 - \alpha_2 O^2 dN_t^2. \end{aligned} \tag{9}$$

The order sizes, O_i^1 and O_i^2, are assumed to be identically independent (i.i.d.) random variables. The assumption of independence between sizes cannot be checked statistically as we don't have information about volumes (the interest rate market being mainly an "over the counter"). However this assumption is common in the literature about micro-structure, as e.g. in Bacry et al. (2013). Their densities are noted $\nu_1(z)$ and $\nu_2(z)$ and their first and second moments are noted $\mu_1 = \mathbb{E}(O^1), \mu_2 = \mathbb{E}(O^2), \eta_1 = \mathbb{E}\left((O^1)^2\right), \eta_2 = \mathbb{E}\left((O^2)^2\right)$. In numerical illustrations, order sizes are exponential random variables. The positiveness of orders ensures the identifiability of the model. Notice however that most of results in this paper can be extended to any other distribution.

At this stage, the interest rate is not explicitly mean reverting and there is no warranty that the short term rate will not diverge at long term to extreme positive or negative values. Such divergences can be avoided (at least at short or medium term) by assuming that arrivals

Table 1
Descriptive statistics, zero coupon rates bootstrapped from swap curves from 3/05/2004 to 25/7/2014. The two last columns contain sample auto-correlation at displacement of 175 days and 250 days of trading.

Maturity	Mean	Std. dev.	Minimum	Maximum	ρ(175 days)	ρ(250 days)
1	0.0217	0.0147	0.0029	0.0545	0.7929	0.6759
2	0.0229	0.0140	0.0031	0.0548	0.8114	0.7315
3	0.0245	0.0134	0.0038	0.0540	0.8222	0.7614
4	0.0261	0.0127	0.0049	0.0528	0.8265	0.7746
5	0.0277	0.0121	0.0062	0.0519	0.8271	0.7800
6	0.0292	0.0115	0.0077	0.0513	0.8273	0.7837
7	0.0305	0.0110	0.0093	0.0510	0.8265	0.7852
8	0.0317	0.0105	0.0109	0.0508	0.8252	0.7855
9	0.0327	0.0102	0.0124	0.0507	0.8232	0.7844
10	0.0337	0.0099	0.0138	0.0509	0.8197	0.7813
12	0.0353	0.0094	0.0161	0.0513	0.8114	0.7728
15	0.0370	0.0090	0.0187	0.0516	0.7990	0.7596
20	0.0380	0.0089	0.0187	0.0526	0.7904	0.7507

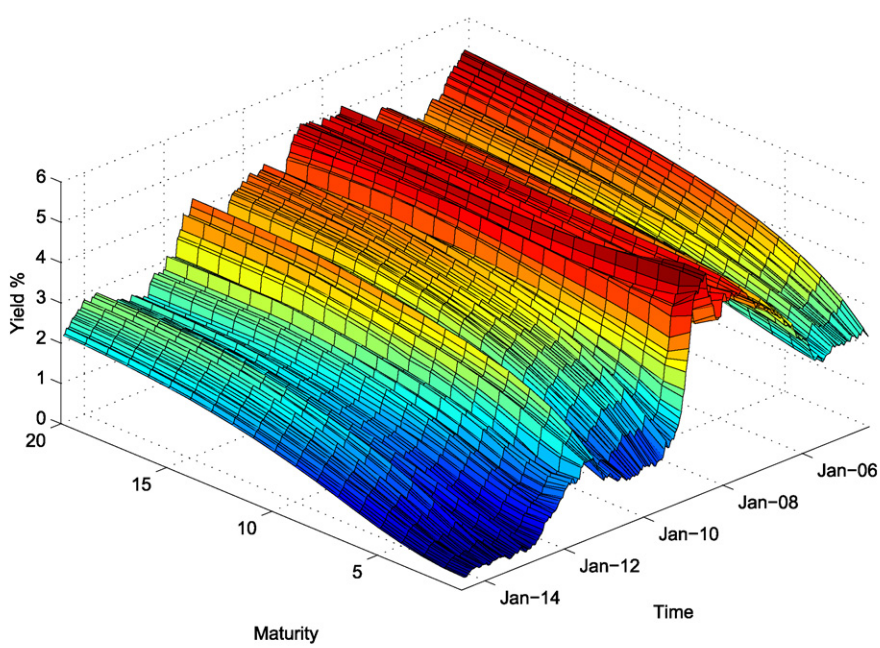

Fig. 2. This graph shows the evolution of zero coupon rates, bootstrapped from swap curves, over a ten year period (3/05/2004 to 25/7/2014).

of bid and ask orders are mutually dependent. If new bid (resp. ask) orders raise the probability of ask (resp. bid) order arrivals, we may expect a stable behavior for X_t. From an economic point of view, this mutual excitation also makes sense. By the end of 2011, most of European countries (e.g. Germany, France or Netherlands) issued debts to answer to the high demand for their national bonds and indirectly benefited from low interest rates to finance their deficit. Mathematically, the mutual and self excitation is obtained by assuming that intensities of order arrivals are random processes governed by the following equations:

$$d\lambda_t^i = \kappa_i \left(c_i - \lambda_t^i\right) dt + \delta_{i,1} dL_t^1 + \delta_{i,2} dL_t^2 \quad i = 1, 2, \tag{10}$$

where $\delta_{i,j}$ for $i, j = 1, 2$, are constant. Coefficients $\delta_{1,2}$ and $\delta_{2,1}$ set the cross impact of demand on supply and vice versa. They measure the dependence between them and can capture some interesting stylized facts like the impact of bond issuance during a period of low interest rates. E.g. if $\delta_{12} > 0$, the frequency of bonds issuance increases when the demand, L_t^2, steps up and drives down interest rates according to Eq. (9). Coefficients $\delta_{1,1}$ and $\delta_{2,2}$ set the self-excitation levels.

Contrary to affine models like the Cox et al. (1985)model, the chosen dynamics for r_t allows negative interest rates but we don't consider it as a limitation. Indeed, since the 2012 crisis of the European sovereign debts, we have observed several periods during which short term rates (sovereign or interbank) were negative (e.g. in 2014, the EONIA was negative 61 times over 254 days of trading). On the other hand, the probability of observing negative rates can be restricted by an appropriate choice for the dynamics of N_t^1, N_t^2, as discussed later in Section 4.

As shown in Errais et al. (2010), if $J_t^i = \left(L_t^i, N_t^i\right)$, the process $(\lambda_t^1, J_t^1, \lambda_t^2, J_t^2)$ is a Markov process in the state space $D = \left(\mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{N}\right)^2$ and its infinitesimal generator for any function $g : D \to \mathbb{R}$ with partial derivatives $g_{\lambda_1}, g_{\lambda_2}$, is such that

$$\begin{aligned}\mathcal{A}g\left(\lambda_t^1, J_t^1, \lambda_t^2, J_t^2\right) &= \kappa_1 \left(c_1 - \lambda_t^1\right) g_{\lambda^1} + \kappa_2 \left(c_2 - \lambda_t^2\right) g_{\lambda^2} \\ &\quad + \lambda_t^1 \int_{-\infty}^{+\infty} g\left(\lambda_t^1 + \delta_{1,1} z, J_t^1 + (z,1)^\top, \lambda_t^2 + \delta_{2,1} z, J_t^2\right) \\ &\quad - g\left(\lambda_t^1, J_t^1, \lambda_t^2, J_t^2\right) d\nu_1(z) \\ &\quad + \lambda_t^2 \int_{-\infty}^{+\infty} g\left(\lambda_t^1 + \delta_{1,2} z, J_t^1, \lambda_t^2 + \delta_{2,2} z, J_t^2 + (z,1)^\top\right) \\ &\quad - g\left(\lambda_t^1, J_t^1, \lambda_t^2, J_t^2\right) d\nu_2(z).\end{aligned} \tag{11}$$

Under mild conditions, the expectation of $g(.)$ is equal to the integral of the expected infinitesimal generator:

$$\begin{aligned}\mathbb{E}\left(g\left(\lambda_T^1, J_T^1, \lambda_T^2, J_T^2\right) | \mathcal{F}_t\right) &= g\left(\lambda_t^1, J_t^1, \lambda_t^2, J_t^2\right) \\ &\quad + \mathbb{E}\left(\int_t^T \mathcal{A}g\left(\lambda_s^1, J_s^1, \lambda_s^2, J_s^2\right) ds | \mathcal{F}_t\right) \\ &= g\left(\lambda_t^1, J_t^1, \lambda_t^2, J_t^2\right) \\ &\quad + \int_t^T \mathbb{E}\left(\mathcal{A}g\left(\lambda_s^1, J_s^1, \lambda_s^2, J_s^2\right) | \mathcal{F}_t\right) ds.\end{aligned} \tag{12}$$

The derivative of this expectation with respect to time is equal to its expected infinitesimal generator:

$$\frac{\partial}{\partial T} \mathbb{E}\left(g\left(\lambda_T^1, J_T^1, \lambda_T^2, J_T^2\right) | \mathcal{F}_t\right) = \mathbb{E}\left(\mathcal{A}g\left(\lambda_T^1, J_T^1, \lambda_T^2, J_T^2\right) | \mathcal{F}_t\right), \tag{13}$$

The next proposition relies on this last feature to calculate the first moments of intensities.

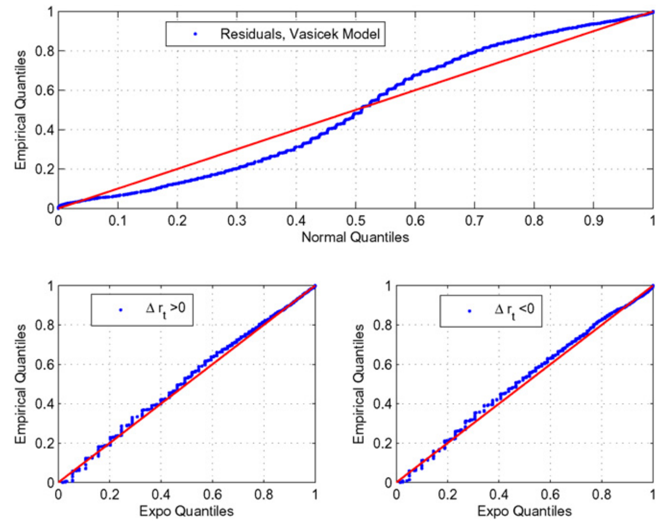

Fig. 3. The first subplot presents the QQ plot of normed residuals for the Vasicek model fitted to the one year swap rate. The second and third subplots display QQ plots of positive and negative variations of interest rates versus exponential distributions of O^1 and O^2.

Proposition 2.1. *Let $m_i(t)$ denote the expected intensity* [1], $\mathbb{E}\left(\lambda_t^i\right)$, *for $i = 1, 2$. They are given by the following expressions*

$$\begin{aligned}\begin{pmatrix} m_1(t) \\ m_2(t) \end{pmatrix} &= V \begin{pmatrix} \frac{1}{\gamma_1}\left(e^{\gamma_1 t} - 1\right) & 0 \\ 0 & \frac{1}{\gamma_2}\left(e^{\gamma_2 t} - 1\right) \end{pmatrix} V^{-1} \begin{pmatrix} \kappa_1 c_1 \\ \kappa_2 c_2 \end{pmatrix} \\ &\quad + V \begin{pmatrix} e^{\gamma_1 t} & 0 \\ 0 & e^{\gamma_2 t} \end{pmatrix} V^{-1} \begin{pmatrix} \lambda_0^1 \\ \lambda_0^2 \end{pmatrix},\end{aligned} \tag{14}$$

where $\gamma_{1,2}$ are constant,

$$\begin{aligned}\gamma_{1,2} :=& \frac{1}{2}\left((\delta_{1,1}\mu_1 - \kappa_1) + (\delta_{2,2}\mu_2 - \kappa_2)\right) \\ & \pm \frac{1}{2}\sqrt{\left((\delta_{1,1}\mu_1 - \kappa_1) - (\delta_{2,2}\mu_2 - \kappa_2)\right)^2 + 4\delta_{1,2}\delta_{2,1}\mu_1\mu_2},\end{aligned} \tag{15}$$

V, V^{-1} are matrices given by

$$V = \begin{pmatrix} -\delta_{1,2}\mu_2 & -\delta_{1,2}\mu_2 \\ (\delta_{1,1}\mu_1 - \kappa_1) - \gamma_1 & (\delta_{1,1}\mu_1 - \kappa_1) - \gamma_2 \end{pmatrix} \tag{16}$$

$$V^{-1} = \frac{1}{\Upsilon} \begin{pmatrix} (\delta_{1,1}\mu_1 - \kappa_1) - \gamma_2 & \delta_{1,2}\mu_2 \\ \gamma_1 - (\delta_{1,1}\mu_1 - \kappa_1) & -\delta_{1,2}\mu_2 \end{pmatrix} \tag{17}$$

and Υ is the determinant of V defined by

$$\Upsilon := -\delta_{1,2}\mu_2\sqrt{\left((\delta_{1,1}\mu_1 - \kappa_1) - (\delta_{2,2}\mu_2 - \kappa_2)\right)^2 + 4\delta_{1,2}\delta_{2,1}\mu_1\mu_2}. \tag{18}$$

The next corollary is an immediate consequence of this last proposition.

Corollary 2.2. *The expectation of X_t is equal to*

$$\mathbb{E}\left(X_t | \mathcal{F}_0\right) = X_0 + \alpha_1\mu_1 \int_0^t m_1(s) ds - \alpha_2\mu_2 \int_0^t m_2(s) ds \tag{19}$$

[1] Remark that $\mathbb{E}\left(.|\mathcal{F}_0\right)$ is abusively denoted by $\mathbb{E}\left(.\right)$ in later developments.

The dynamics of X_t is not mean-reverting and the short term rate, equal to $\varphi(t) + X_t$, may diverge at long term if $\varphi(t)$ is constant. Even if numerical illustrations confirm that the mechanism of mutual excitation limits this phenomenon for short and medium maturities (up to 20 years), divergence at long term can be controlled by choosing an adapted $\varphi(t)$. For example, if we want to keep constant the average short term rate, we can set $\varphi(t)$ to $r_0 - \mathbb{E}(X_t|\mathcal{F}_0)$, such as defined by Eq. (6). In this case, the expected rate stays equal to $\mathbb{E}(r_t|\mathcal{F}_0) = r_0$, whatever the maturity. If we want that expected short term rates revert at long term to r_∞, $\varphi(t)$ can be set to $(e^{-at}-1)r_\infty + e^{-at}r_0 - \mathbb{E}(X_t|\mathcal{F}_0)$, where a is the speed of mean reversion. The next system of ODEs that rules variances and correlation of intensities is provided in the next proposition.

Proposition 2.3. *Let us denote the variance of λ_i by $V_i(t) = \mathbb{E}\left(\left(\lambda_t^i\right)^2\right) - (m_i(t))^2$ for $i = 1, 2$ and their covariance by $V_3(t) = \mathbb{E}(\lambda_t^1\lambda_t^2) - m_1(t)m_2(t)$. They are solutions of the following system of ODEs:*

$$\frac{\partial}{\partial t}\begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix} = \begin{pmatrix} \delta_{1,1}^2\eta_1 & \delta_{1,2}^2\eta_2 \\ \delta_{2,1}^2\eta_1 & \delta_{2,2}^2\eta_2 \\ \delta_{1,1}\delta_{2,1}\eta_1 & \delta_{2,2}\delta_{1,2}\eta_2 \end{pmatrix}\begin{pmatrix} m_1(t) \\ m_2(t) \end{pmatrix}$$
$$+\begin{pmatrix} 2(\delta_{1,1}\mu_1 - \kappa_1) & 0 & 2\delta_{1,2}\mu_2 \\ 0 & 2(\delta_{2,2}\mu_2 - \kappa_2) & 2\delta_{2,1}\mu_1 \\ \delta_{2,1}\mu_1 & \delta_{1,2}\mu_2 & (\delta_{1,1}\mu_1 + \delta_{2,2}\mu_2) - \kappa_1 - \kappa_2 \end{pmatrix}$$
$$\times \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix}, \qquad (20)$$

with initial conditions

$$V_i(0) = 0 \qquad \text{for } i = 1, 2, 3.$$

By construction, the model behaves like a Brownian motion with a random volatility. In this sense, the model is close to the GARCH model of Engle (1982). Jumps in our model, by virtue of their self- and cross-excitation, introduce a feedback element. As mentioned by Aït-Sahalia et al. (2015), this aspect of the model can be thought of as playing the same role for jumps as GARCH does for volatility. Namely, the GARCH model, applied to interest rates, introduces feedback from variation of rates to volatility and back: large deviations lead to large volatilities which then make it more likely to observe large deviations. In the absence of further excitation, volatility then reverts to its steady state level. Here, similarly, jumps lead to larger jump intensities, which then make it more likely to observe further jumps. In the absence of further excitation, jump intensities then revert to their steady state level. For this reason, we compare in the numerical application our model with the GARCH one.

Table 2
This table contains the parameters defining dX_t under the real measure and their standard errors.

Parameter	Value	Std. err.	Parameter	Value	Std. err.
κ_1	3.86	0.0187	$\delta_{1,1}$	5619.40	10.3072
κ_2	3.93	0.0196	$\delta_{1,2}$	−4681.94	8.1994
c_1	120.00	0.0643	$\delta_{2,1}$	−5486.30	10.1354
c_2	134.63	0.0621	$\delta_{2,2}$	4608.52	7.5390
α_1	1		α_2	1	
ρ_1	5653.08	6.0047	ρ_2	5576.20	6.0875

The next proposition presents the moment generating function of X_t and of its integral. This result is used later to infer the price of a bond and its dynamics under an equivalent measure.

Proposition 2.4. *Let $\psi_1(.)$ and $\psi_2(.)$ denote the moment generating functions of O^1 and O^2:*

$$\psi_i(w) := \mathbb{E}\left(e^{wO^i}\right) \quad i = 1, 2. \qquad (21)$$

The moment generating function of $w_0X_T - w_1\int_t^T X_s ds + \begin{pmatrix} w_2 \\ w_3 \end{pmatrix}^\top \begin{pmatrix} \lambda_T^1 \\ \lambda_T^2 \end{pmatrix}$ is an affine function of X_t and of intensities

$$\mathbb{E}\left(e^{w_0X_T - w_1\int_t^T X_s ds + \begin{pmatrix} w_2 \\ w_3 \end{pmatrix}^\top \begin{pmatrix} \lambda_T^1 \\ \lambda_T^2 \end{pmatrix}} \mid \mathcal{F}_t\right)$$
$$= \exp\left((w_0 - w_1(T-t))X_t + A(t,T) + \begin{pmatrix} B_1(t,T) \\ B_2(t,T) \end{pmatrix}^\top \begin{pmatrix} \lambda_t^1 \\ \lambda_t^2 \end{pmatrix}\right) \qquad (22)$$

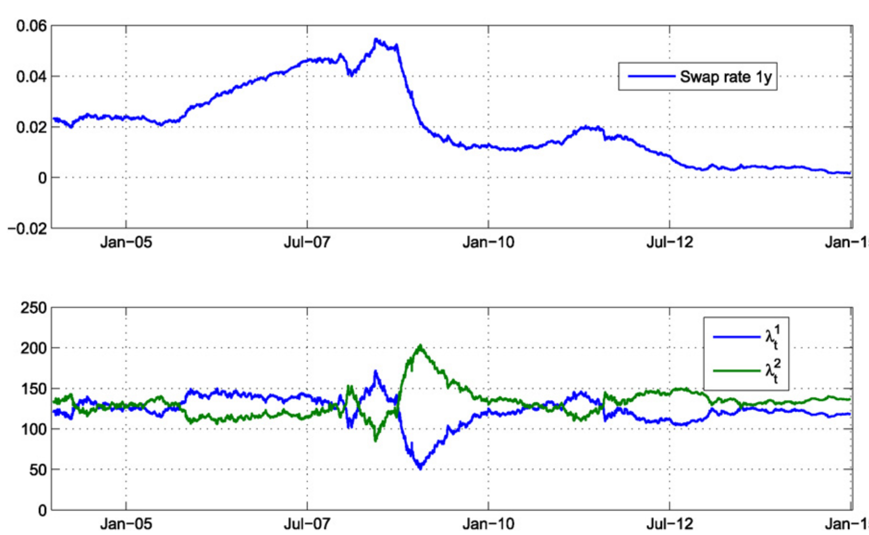

Fig. 4. The first subplot presents the history of the one year swap rate, to which the model is fitted. The second graph shows the sample paths of λ_t^1 and λ_t^2 obtained by log likelihood maximization.

where $A(t,T)$, $B_1(t,T)$ and $B_2(t,T)$ are solutions of a system of ODEs as follows:

$$\begin{cases} \frac{\partial}{\partial t}B_1(t,T) = \kappa_1 B_1(t,T) - [\psi_1\left(B_1(t,T)\delta_{1,1} + w_0\alpha_1 - w_1\alpha_1(T-t)\right. \\ \qquad\qquad \left.+B_2(t,T)\delta_{2,1}\right) - 1] \\ \frac{\partial}{\partial t}B_2(t,T) = \kappa_2 B_2(t,T) - [\psi_2\left(B_1(t,T)\delta_{1,2} - w_0\alpha_2 + w_1\alpha_2(T-t)\right. \\ \qquad\qquad \left.+B_2(t,T)\delta_{2,2}\right) - 1] \\ \frac{\partial}{\partial t}A(t,T) = -\kappa_1 c_1 B_1(t,T) - \kappa_2 c_2 B_2(t,T) \end{cases} \tag{23}$$

with the terminal conditions $A(T,T)=0$, $B_1(T,T)=w_2$, $B_2(T,T)=w_3$.

Notice that for sufficiently large $\omega_{i=1,\ldots,4}$, there is no warranty that the expectation (Eq. (22)) is finite. We will see in Section 4 that it is however possible to compute the probability density function of X_t, by inverting the moment generating function (22), with the discrete Fourier transform of Proposition 3.8.

3. Equivalent exponential affine measures and bond pricing.

As the moment generating function of X_t is an affine function of $(\lambda_t^1, \lambda_t^2, L_t^1, L_t^2)$, we study exponential affine changes of measure and show that the dynamics of interest rates is preserved under the new measure. These equivalent measures are induced by an exponential martingale of the form

$$\begin{aligned} M_t(\theta_1,\theta_2) := \exp\Bigg(& (a_1(\theta_1,\theta_2),\ a_2(\theta_1,\theta_2)) \begin{pmatrix} \lambda_t^1 \\ \lambda_t^2 \end{pmatrix} \\ & +(\theta_1,\theta_2)\begin{pmatrix} L_t^1 \\ L_t^2 \end{pmatrix} - \varphi(\theta_1,\theta_2)t \Bigg) \end{aligned} \tag{24}$$

where $\theta_1,\ \theta_2\ \in\ \mathbb{R}$ and are assimilated later to risk premiums. Zhang et al. (2009) use a similar change of measure to simulate rare events, of a one dimension Hawkes process but with constant jumps. In our framework, jumps are random and the affine change of measure modifies both frequencies and distributions of jumps. Before detailing this point, the next proposition introduces the necessary conditions that (θ_1,θ_2) fulfill to guarantee that $M_t(\theta_1,\theta_2)$ is a local martingale.

Proposition 3.1. *If for any given couple of parameters (θ_1,θ_2), there exist suitable solutions $a_1(\theta_1,\theta_2)$ and $a_2(\theta_1,\theta_2)$ for the system of equations*

$$\begin{cases} a_1(\theta_1,\theta_2)\kappa_1 - (\psi_1(a_1(\theta_1,\theta_2)\delta_{1,1} + a_2(\theta_1,\theta_2)\delta_{2,1} + \theta_1) - 1) = 0 \\ a_2(\theta_1,\theta_2)\kappa_2 - (\psi_2(a_2(\theta_1,\theta_2)\delta_{2,2} + a_1(\theta_1,\theta_2)\delta_{1,2} + \theta_2) - 1) = 0 \end{cases} \tag{25}$$

where $\psi_i(w) = \mathbb{E}\left(e^{wO^i}\right)$ for $i = 1,2$, and if $\varphi(\theta_1,\theta_2)$ is a linear combination of these solutions

$$\varphi(\theta_1,\theta_2) = a_1(\theta_1,\theta_2)\kappa_1 c_1 + a_2(\theta_1,\theta_2)\kappa_2 c_2 \tag{26}$$

then $M_t(\theta_1,\theta_2)$ is a local martingale.

Assuming the existence of suitable solutions for the system (Eq. (25)), an equivalent measure Q^{θ_1,θ_2} is defined by

$$\left.\frac{dQ^{\theta_1,\theta_2}}{dP}\right|_{\mathcal{F}_t} = \frac{M_t(\theta_1,\theta_2)}{M_0(\theta_1,\theta_2)} \tag{27}$$

and may be used as risk neutral measure by investors. In this case, the dynamics of intensities and aggregate supply or demand is modified but is still a bivariate Hawkes process:

Proposition 3.2. *Let $N_t^{1,Q}$ and $N_t^{2,Q}$ be counting processes with respective intensities*

$$\begin{cases} \lambda_t^{1,Q} = \mathbb{E}\left(e^{(a_1\delta_{1,1}+a_2\delta_{2,1}+\theta_1)O^1}\right)\lambda_t^1 \\ \lambda_t^{2,Q} = \mathbb{E}\left(e^{(a_2\delta_{2,2}+a_1\delta_{1,2}+\theta_2)O^2}\right)\lambda_t^2 \end{cases}, \tag{28}$$

under the equivalent measure Q^{θ_1,θ_2}. On the other hand, if $O^{1,Q}$, $O^{2,Q}$ denote random variables defined by the following moment generating functions

$$\begin{cases} \psi_1^Q(z) := \mathbb{E}\left(e^{zO^{1,Q}}\right) & = \frac{\psi_1(z+(\delta_{1,1}a_1+\delta_{2,1}a_2+\theta_1))}{\psi_1(a_1\delta_{1,1}+a_2\delta_{2,1}+\theta_1)} \\ \psi_2^Q(z) := \mathbb{E}\left(e^{zO^{2,Q}}\right) & = \frac{\psi_2(z+(a_2\delta_{2,2}+a_1\delta_{1,2}+\theta_2))}{\psi_2(a_2\delta_{2,2}+a_1\delta_{1,2}+\theta_2)} \end{cases} \tag{29}$$

and if $L_t^{1,Q}$, $L_t^{2,Q}$ are defined by the jump processes

$$L_t^{i,Q} = \sum_{k=1}^{N_t^{i,Q}} O_k^{i,Q} \quad i = 1,2, \tag{30}$$

then intensities λ_t^j are driven by the following SDE under Q^{θ_1,θ_2}

$$d\lambda_t^i = \kappa_i\left(c_i - \lambda_t^i\right)dt + \delta_{i,1}dL_t^{1,Q} + \delta_{i,2}dL_t^{2,Q} \quad i = 1,2. \tag{31}$$

In numerical applications, sizes of orders under P are exponential random variables and their probability density functions is defined by two parameters $\rho_1, \rho_2 \in \mathbb{R}^+$ as follows:

$$\nu_1(z) = \rho_1 e^{-\rho_1 z}1_{\{z\geq 0\}} \quad \nu_2(z) = \rho_2 e^{\rho_2 z}1_{\{z\leq 0\}}. \tag{32}$$

Table 3
Deviances and p-values testing the presence of mutual and self excitation. All hypotheses are rejected with a high confidence rate.

Hypothesis	Deviance	p-Value
$\delta_{1,1}=0$	24.08	0.0001%
$\delta_{1,2}=0$	44.16	0.0000%
$\delta_{2,1}=0$	24.05	0.0001%
$\delta_{2,2}=0$	202.88	0.0000%
$\delta_{2,1}=\delta_{1,2}=0$	68.21	0.0000%
$\delta_{1,1}=\delta_{2,2}=0$	290.24	0.0000%

Table 4
Parameters of the Vasicek model, obtained by log likelihood maximization.

Parameters	Estimates	Std. error
a	0.014	0.0003
b	−0.103	0.0017
σ	0.004	0.0001
Loglik.	19078	
AIC	−38150	

Table 5
Parameters of a GARCH (1,0) and GARCH (2,0) model for r_t.

Garch(1,0)	Estimate	Std. error	t value	Pr(> |t|)
μ	1.652e − 06	4.543e − 06	0.364	0.716
α_0	4.844e − 08	1.716e − 09	28.222	< 2e − 16 ***
β_1	4.211e − 01	4.006e − 02	10.512	< 2e − 16 ***
Loglik.	19,311			

In this case, first and second moments of O_1 and O_2 are respectively equal to $\mu_1 = \frac{1}{\rho_1}$, $\mu_2 = -\frac{1}{\rho_2}$ and to $\eta_i = \frac{2}{(\rho_i)^2}$. The moment generating functions are given by $\psi_1(z) = \frac{\rho_1}{\rho_1 - z}$ for $z < \rho_1$ and $\psi_2(z) = \frac{\rho_2}{\rho_2 + z}$ for $z > -\rho_2$. In this particular, we have the following interesting corollary:

Corollary 3.3. *The distribution of orders are exponential under P and Q and the densities, noted $\nu_i^Q(z)$ under Q, are defined by the following parameters:*

$$\rho_1^Q = \rho_1 - (\delta_{1,1} a_1 + \delta_{2,1} a_2 + \theta_1)$$
$$\rho_2^Q = \rho_2 + (a_2 \delta_{2,2} + a_1 \delta_{1,2} + \theta_2)\,.$$

Under Q^{θ_1,θ_2}, dynamics of intensities are preserved as shown in the next corollary.

Corollary 3.4. *Intensities of counting processes $N_t^{1,Q}$ and $N_t^{2,Q}$ are Hawkes processes having the same structure under Q as these under the real measure P*

$$d\lambda_t^{i,Q} = \kappa_i \left(c_i^Q - \lambda_t^{i,Q}\right) dt + \delta_{i,1}^Q dL_t^{1,Q} + \delta_{i,2}^Q dL_t^{2,Q} \quad i = 1,2. \tag{33}$$

where the parameters defining the process under Q are

$$c_1^Q = c_1 \psi_1(a_1 \delta_{1,1} + a_2 \delta_{2,1} + \theta_1)$$
$$c_2^Q = c_2 \psi_2(a_2 \delta_{2,2} + a_1 \delta_{1,2} + \theta_2)$$
$$\delta_{1,j}^Q = \delta_{1,j} \psi_1(a_1 \delta_{1,1} + a_2 \delta_{2,1} + \theta_1) \quad j = 1,2$$
$$\delta_{2,j}^Q = \delta_{2,j} \psi_2(a_2 \delta_{2,2} + a_1 \delta_{1,2} + \theta_2) \quad j = 1,2.$$

This corollary is proved by combining Eqs. (28) and (31). If markets participants adopt an equivalent exponential affine measure for the risk neutral one, the price of a zero coupon bond is equal to the expected discount factor, under this risk neutral measure. The price is denoted by

$$\begin{aligned} P(t,T,\lambda_t^1,J_t^1,\lambda_t^2,J_t^2) &= \mathbb{E}^Q\left(e^{-\int_t^T r_s ds} \mid \mathcal{F}_t\right) \\ &= e^{-\int_t^T \varphi(s) ds} \mathbb{E}^Q\left(e^{-\int_t^T X_s ds} \mid \mathcal{F}_t\right) \end{aligned} \tag{34}$$

and the expectation in the left term of the bond price is provided in the corollary that follows, that is proved by combining Proposition 2.4 with Corollary 3.4.

Corollary 3.5.

$$\mathbb{E}^Q\left(e^{-\int_t^T X_s ds} \mid \mathcal{F}_t\right) = \exp\left(-X_t(T-t) + A(t,T) + \begin{pmatrix} B_1(t,T) \\ B_2(t,T) \end{pmatrix}^\top \begin{pmatrix} \lambda_t^{1Q} \\ \lambda_t^{2Q} \end{pmatrix}\right) \tag{35}$$

where $A(t,T)$, $B_1(t,T)$ and $B_2(t,T)$ are solutions of a system of ODEs as follows:

$$\begin{cases} \frac{\partial}{\partial t} B_1(t,T) = \kappa_1 B_1(t,T) \\ \quad -\left[\psi_1^Q\left(B_1(t,T)\delta_{1,1}^Q - \alpha_1(T-t) + B_2(t,T)\delta_{2,1}^Q\right) - 1\right] \\ \frac{\partial}{\partial t} B_2(t,T) = \kappa_2 B_2(t,T) \\ \quad -\left[\psi_2^Q\left(B_1(t,T)\delta_{1,2}^Q + \alpha_2(T-t) + B_2(t,T)\delta_{2,2}^Q\right) - 1\right] \\ \frac{\partial}{\partial t} A(t,T) = -\kappa_1 c_1^Q B_1(t,T) - \kappa_2 c_2^Q B_2(t,T) \end{cases} \tag{36}$$

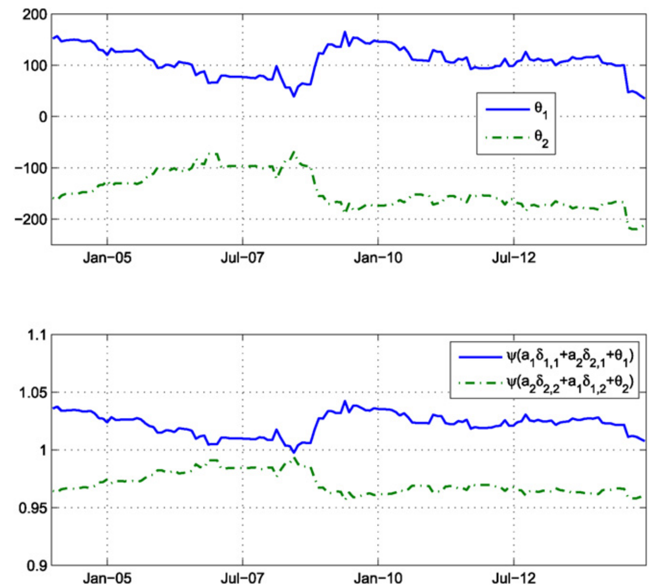

Fig. 5. The first subplot exhibits the history of risk premiums, appraised at regular interval of five days. The second graph shows coefficients $\psi_1(a_1\delta_{1,1} + a_2\delta_{2,1} + \theta_1)$ and $\psi_2(a_2\delta_{2,2} + a_1\delta_{1,2} + \theta_2)$ that multiply the parameters defining λ_t^1 and λ_t^2 under the risk neutral measure.

with the terminal conditions $A(T,T) = 0$, $B_1(T,T) = 0$, $B_2(T,T) = 0$.

The dynamics of bond prices depends upon the random measures of jump processes, noted $L^{1Q}(dt,dz)$ and $L^{2Q}(dt,dz)$ and such that

$$L_t^{kQ} = \int_0^\infty \int_{-\infty}^\infty L^{kQ}(dt,dz) \quad k = 1,2.$$

Furthermore, the expectation of these measures are equal to $\mathbb{E}^Q\left(L^{kQ}(dt,dz)|\mathcal{F}_t\right) = \lambda_t^{kQ}\nu_k(z)\,dz\,dt$, for $k = 1,2$. The next corollary details the infinitesimal dynamics of bond prices:

Corollary 3.6. *Bond prices, $P\left(t,T,\lambda_t^{1Q},J_t^{1Q},\lambda_t^{2Q},J_t^{2Q}\right)$, are ruled by the following SDE:*

$$\begin{aligned} dP &= P\,r_t\,dt - \lambda_t^{1Q} P\left[\psi_1^Q\left(B_1(t,T)\delta_{1,1}^Q - \alpha_1(T-t) + B_2(t,T)\delta_{2,1}^Q\right) - 1\right]dt \\ &\quad + P\int_{-\infty}^{+\infty}\left[\exp\left(\left(B_1(t,T)\delta_{1,1}^Q - \alpha_1(T-t) + B_1(t,T)\delta_{2,1}^Q\right)z\right) - 1\right] \\ &\quad \times L^{1Q}(dt,dz) \\ &\quad - \lambda_t^{2Q} P\left[\psi_2^Q\left(B_1(t,T)\delta_{1,2}^Q + \alpha_2(T-t) + B_2(t,T)\delta_{2,2}^Q\right) - 1\right]dt \\ &\quad + P\int_{-\infty}^{+\infty}\left[\exp\left(\left(B_1(t,T)\delta_{1,2}^Q + \alpha_2(T-t) + B_2(t,T)\delta_{2,2}^Q\right)z\right) - 1\right] \\ &\quad \times L^{2Q}(dt,dz) \end{aligned} \tag{37}$$

where $L^{1Q}(dt,dz)$ and $L^{2Q}(dt,dz)$ are the random measures of jump processes.

From the last corollary, we infer that the expected growth rate for the bond price under Q is well equal to the short-term rate, $\mathbb{E}^Q\left(\frac{dP}{P}|\mathcal{F}_t\right) = r_t dt$, as the sum of all other terms in Eq. (37) is a martingale.

3.1. Pricing of options

This section illustrates how the model is used for the pricing of interest rate derivatives, under a forward measure. The yield of maturity $T-S$, at time T is denoted by $Y(T,S)$ and is defined by

$$\begin{aligned} Y(T,S) &:= -\frac{1}{S-T}\log P(T,S) \\ &= X_T + \frac{1}{S-T}\left(\int_T^S \varphi(s)ds - A(T,S)\right) - \frac{1}{S-T}\begin{pmatrix} B_1(T,S) \\ B_2(T,S)\end{pmatrix}^{\top}\begin{pmatrix}\lambda_T^{1Q} \\ \lambda_T^{2Q}\end{pmatrix} \end{aligned} \tag{38}$$

On the other hand, the payoff paid at time $S \geq T$ by an European option written on $Y(T,S)$ is denoted by $V(Y(T,S))$. Examples of such instruments are: caplets ($V(Y(T,S)) = N(S-T)[Y(T,S)-k]_+$), floorlets ($V(Y(T,S)) = N(S-T)N[k-Y(T,S)]_+$) or options of zero coupon bonds ($V(Y(T,S)) = N[\exp(-Y(T,S)(S-T))-k]_+$), where N and k are respectively the principal and the strike. The option price is the expectation of this discounted payoff under the risk neutral measure.

$$f(t, r_t, \lambda_t) = \mathbb{E}^Q\left(e^{-\int_t^S r_s ds} V(Y(T,S)) \mid \mathcal{F}_t\right) \tag{39}$$

As recommended by Brigo and Mercurio (2007), it is better to evaluate this last expression under the S-forward measure. This avoids numerical inaccuracies related to the approximation of $\exp\left(-\int_t^S r_s ds\right)$, because the discount factor is drawn out of Eq. (39), under the forward measure. If the market admits at least one risk neutral measure Q, an equivalent probability measures to Q is defined by the technique of changes of numeraire. The S-forward measure has as numeraire, the zero coupon bond of maturity S. Under this measure, the price of any financial assets, divided by the numeraire $P(t,S)$, is a martingale and the price of the derivative is equal to

$$\begin{aligned} \mathbb{E}^Q\left(e^{-\int_t^S r_s ds} V(Y(T,S)) \mid \mathcal{F}_t\right) &= P(t,S)\mathbb{E}^S\left(V(Y(T,S)) \mid \mathcal{F}_t\right) \\ &= P(t,S)\int_0^{+\infty} V(y) f_{Y(T,S)}(y) dy \end{aligned}$$

where $f_{Y(T,S)}(y)$ is the density of $Y(T,S)$ under the forward measure. If $B(t)$ points out here the market value of a cash account, $B_t = e^{\int_0^t r_s ds}$, the Radon Nykodym derivative defining the S-forward measure, is equal to

$$\frac{dF^S}{dQ} = \frac{1}{B_S}\frac{B_0}{P(0,S)} = \left(e^{\int_0^S r_s ds}\mathbb{E}^Q\left(e^{-\int_0^S r_s ds} | \mathcal{F}_0\right)\right)^{-1}$$

To calculate the expected payoff under F^S, the easiest approach consists to approximate the probability density function of $Y(T,S)$ by a discrete Fourier transform. To perform a such calculation, the moment generating function of the yield is needed.

Table 6
This table presents the parameters defining the dynamics of r_t under the risk neutral measure, on the 28/11/2014. They are obtained by adjusting historical parameters according to Corollary 3.4. The values of κ_1, κ_2, c_1^Q and c_2^Q in the column "Best fit" are these for which the model replicates accurately the yield curve. But they cannot be reconciled with historical parameters by an affine change of measure.

	Value under Q	Best fit		Value under Q
κ_1	3.863	1.4463	$\delta_{1,1}^Q$	5662.25
κ_2	3.928	7.6587	$\delta_{1,2}^Q$	−4717.21
c_1^Q	120.91	112.54	$\delta_{2,1}^Q$	−5270.10
c_2^Q	129.33	128.78	$\delta_{2,2}^Q$	4426.65
$\lambda_0^{1,Q}$	119.08		$\lambda_0^{1,Q}$	131.08
r_0	0.0018			
ρ_1^Q	5609.62		ρ_2^Q	5804.41

Corollary 3.7. *The moment generating function of $Y(T,S)$ at time $t \leq T$ under the forward measure F^S, denoted by $\varphi^{t,S}(w)$, is given by:*

$$\begin{aligned} \varphi^{t,S}(w) = \mathbb{E}^S\left(e^{wY(T,S)} \mid \mathcal{F}_t\right) &= \exp\left(\left(\frac{w}{S-T}\right)\int_T^S \varphi(s)ds + wX_t\right) \\ &\times \exp\left(A^T(t,T) - A^S(t,S) + \left(\begin{pmatrix} B_1^T(t,T) - B_1^S(t,S) \\ B_2^T(t,T) - B_2^S(t,S)\end{pmatrix}^{\top}\begin{pmatrix}\lambda_T^{1Q} \\ \lambda_T^{2Q}\end{pmatrix}\right)\right) \end{aligned}$$

where $A^S(t,S)$, $B_1^S(t,S)$ and $B_2^S(t,S)$ are solutions of the system of ODEs (Eq. (36)) with a maturity S and where $A^T(t,T)$, $B_1^T(t,T)$ and $B_2^T(t,T)$ are solutions of the following system of ODEs:

$$\begin{cases} \frac{\partial}{\partial t}B_1^T(t,T) = \kappa_1 B_1^T(t,T) \\ \quad -\left[\psi_1\left(B_1^T(t,T)\delta_{1,1}^Q + (w-(S-t))\,\alpha_1 + B_2^T(t,T)\delta_{2,1}^Q\right) - 1\right] \\ \frac{\partial}{\partial t}B_2^T(t,T) = \kappa_2 B_2^T(t,T) \\ \quad -\left[\psi_2\left(B_1^T(t,T)\delta_{1,2}^Q - (w-(S-t))\,\alpha_2 + B_2^T(t,T)\delta_{2,2}^Q\right) - 1\right] \\ \frac{\partial}{\partial t}A^T(t,T) = -\kappa_1 c_1^Q B_1^T(t,T) - \kappa_2 c_2^Q B_2^T(t,T) \end{cases} \tag{40}$$

with the terminal conditions $A^T(T,T) = \left(1-\frac{w}{S-T}\right)A^S(T,S)$, $B_1^T(T,T) = \left(1-\frac{w}{S-T}\right)B_1^S(T,S)$, $B_2^T(T,T) = \left(1-\frac{w}{S-T}\right)B_2^S(T,S)$.

The next result introduces the discretization framework to build the density of $Y(T,S)$, under the forward measure. Note that it is possible to use the same algorithm to approach the distribution of r_t under the real and risk neutral measure.

Proposition 3.8. *Let M be the number of steps used in the discrete Fourier transform (DFT) and $\Delta_y = \frac{2y_{max}}{M-1}$ be this step of discretization. Let us denote $\Delta_z = \frac{2\pi}{M\,\Delta_y}$ and*

$$z_j = (j-1)\Delta_z,$$

for $j = 1 \ldots M$. The values of $f_{Y(T,S)}(.)$ at points $y_k = -\frac{M}{2}\Delta_y + (k-1)\Delta_y$ are approached by the sum

$$f_{Y(T,S)}(y_k) \approx \frac{2}{M\,\Delta_y} Re\left(\sum_{j=1}^{M} \delta_j \varphi^{t,S}\left(i\, z_j, r_t, \lambda_t\right)(-1)^{j-1} e^{-i\frac{2\pi}{M}(j-1)(k-1)}\right) \tag{41}$$

where $\delta_j = \frac{1}{2}1_{\{j=1\}} + 1_{\{j\neq 1\}}$.

Once that the density of $Y(T,S)$ is obtained by the discrete Fourier transform, the option price is approached by a weighted sum of payoffs as follows:

$$\mathbb{E}^Q\left(e^{-\int_t^T r_s ds} V(Y(T,S)) \mid \mathcal{F}_t\right) = P(t,T)\sum_{k=1}^{M+1} V(y_k) f_{Y(T,S)}(y_k)\Delta_y \,.$$

The feasibility of this method is illustrated for caplets, in the numerical application.

4. Calibration and numerical applications

To demonstrate that the model is adequate for interest rate modeling, we first perform an econometric calibration. The data set

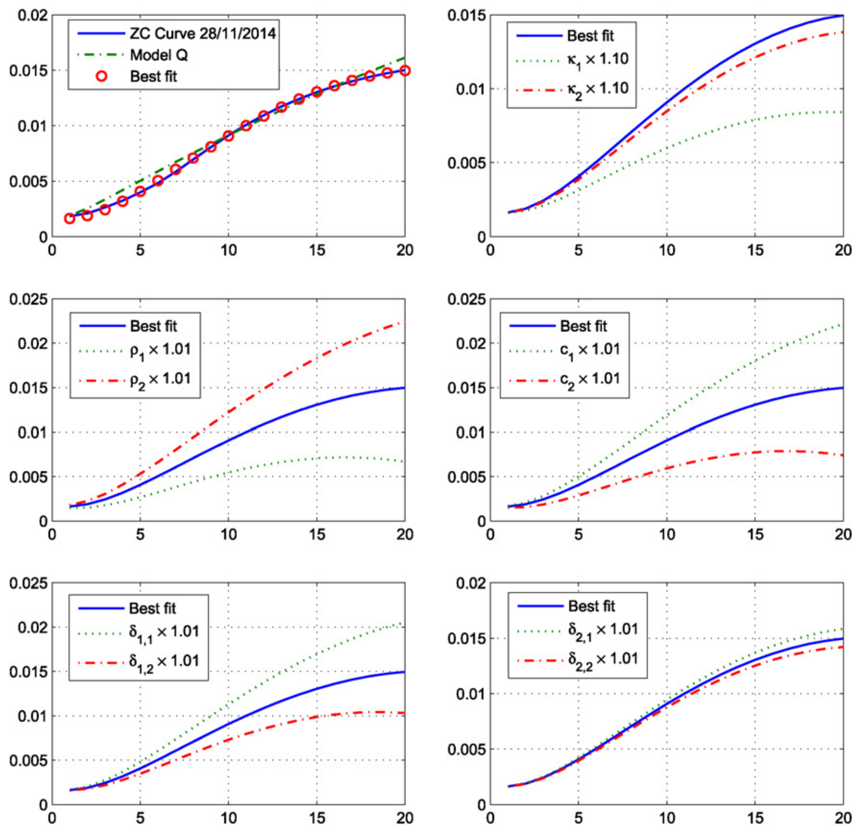

Fig. 6. The first subplot shows the zero coupon yield curve, on the 28/11/2014 and the curves built with the model and sets of parameters of Table 6. The five last subplots illustrate the sensitivity of the yield curve produced by the model to changes of parameters.

used is made up zero coupon rates, bootstrapped from daily Euro swap rates (bid-ask average), observed over ten years (3/05/2004 to 30/12/2014). Swaps are liquid instruments, and their rates are representative of yields of AA corporate bonds. The maturities of considered swaps are running from 1 to 10 years, 12 15 and 20 years. The Bloomberg tickers are EUSA1 to EUSA10, EUSA12, EUSA15 and EUSA20 and the field is PX_LAST. Fig. 2 provides a three-dimensional plot of zero coupon rates. The large amount of temporal variation in the level is visually apparent. The flattening of the curvature during the 2008 crisis, is also clearly visible. Table 1 shows descriptive statistics for swap rates. The typical curve is upward sloping and long term rates are less volatile and more persistent than short term rates (in the sense that their long term auto-correlation is higher).

The parameters that define the dynamics of the short term rate under the real measure, are fitted to the time series of one year swap rates (presented in the first subplot of Fig. 4). Positive and negative jumps are both assumed to be exponential random variables, with means $\frac{1}{\rho_1}$ and $-\frac{1}{\rho_2}$. For this choice of distributions, parameters α_1 and α_2 are redundant and set to one. The function of time $\varphi(t)$ is assumed constant and equal to the one year swap rate, on the 3/05/2004. In practice, $\varphi(t)$ is used to perfectly duplicates the most recent yield curve and to exclude any possibilities of arbitrage in the pricing of interest rate derivatives. As our purpose is here econometric, setting $\varphi(t)$ to a constant is not penalizing. Positive (1360 observations) and negative (1460 observations) variations of the one year rate are respectively assimilated to an increase of supply and increase of demand, the parameters ρ_1 and ρ_2 are adjusted by matching the first moments. The second and third subplots of Fig. 3 confirm that exponential distributions fit well variations of interest rates. The intensities λ_t^1 and λ_t^2 are fitted separately by direct log-likelihood maximization procedures. If daily variations of interest rates are denoted by $\Delta r_i = r_{t_i} - r_{t_{i-1}}$ for $i = 1$ to $n = 2820$ observations and Δ_t is the length of the time interval, the following two optimization problems are solved numerically to find an estimate of parameters:

$$\begin{cases} (\kappa_1, c_1, \delta_{1,1}, \delta_{1,2}) = \arg\max \sum_{i=1}^{n} \log \left(\lambda_{t_i}^1 \Delta_t 1_{\{\Delta r_i > 0\}} + \left(1 - \lambda_{t_i}^2 \Delta_t\right) 1_{\{\Delta r_i \le 0\}} \right) \\ (\kappa_2, c_2, \delta_{2,1}, \delta_{2,2}) = \arg\max \sum_{i=1}^{n} \log \left(\lambda_{t_i}^2 \Delta_t 1_{\{\Delta r_i < 0\}} + \left(1 - \lambda_{t_i}^2 \Delta_t\right) 1_{\{\Delta r_i \ge 0\}} \right) \end{cases} \tag{42}$$

where the intensity of the arrival of jumps is discretized as follows:

$$\lambda_{t_i}^k = \lambda_{t_{i-1}}^k + \kappa_k (c_k - \lambda_{t_{i-1}}) \Delta_t + \delta_{k,1} |\Delta r_i| 1_{\Delta r_i \ge 0} + \delta_{k,2} |\Delta r_i| 1_{\Delta r_i \ge 0}$$
$$k = 1, 2\ ,\ i = 1, \ldots, n.$$

The results of the calibration procedure are presented in Table 2. Notice that given the limited number of parameters and that the calibration is done in three separate steps, the fit does not present any difficulty and is numerically stable. The speeds of mean reversion for the intensities of supply and demand are close and around 3.90. The $\delta_{1,1}$ and $\delta_{2,2}$ measure the level of self excitation and are positive. This confirms the presence of marginal clustering effects in the frequency of orders. The marginal effect of the demand on supply, such as measured by $\delta_{1,2}$, is negative. This means that an upward shift in demand decreases the frequency of supply orders. But as $|\delta_{1,2}| \frac{1}{\rho_2} < \delta_{1,1} \frac{1}{\rho_1}$, this effect is less significant on average than the self excitation. $\delta_{2,1}$ is also negative and then an increase of supply decreases the frequency of demand orders. Compared to the self excitation, this effect is predominant on average as $|\delta_{2,1}| \frac{1}{\rho_1} > \delta_{2,2} \frac{1}{\rho_2}$.

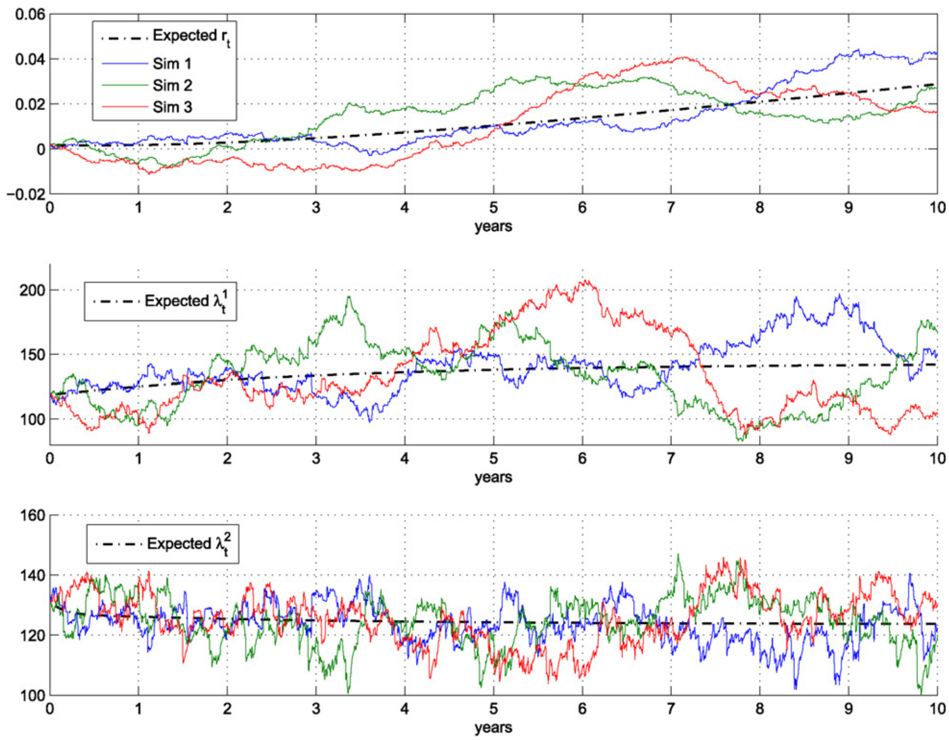

Fig. 7. This graph displays three simulated sample paths of r_t, and the intensity of arrivals of sell (λ_t^1)/buy (λ_t^2) orders. The period is 10 years.

This means that the supply drives the demand for bonds rather than the opposite.

So as to validate whether or not contagion occurs, we test the joint hypothesis that all the coefficients of mutual excitation $\delta_{i,j}$'s are 0. As in Aït-Sahalia et al. (2015), we use a likelihood ratio test based on the deviance. The deviance, noted D, is defined by

$$D = -2 \times \left(\log(Likelihood_{\text{Full model}}) - \log(Likelihood_{\text{reduced model}})\right).$$

According to the results of Wilks (1938), it is asymptotically distributed as a $\mathcal{X}_d^2$ random variable with degrees of freedom equal to the difference between the number of parameters in the full and reduced models. The log likelihood chosen for this test is the sum of maximized terms in Eq. (42). Deviances and their p-values presented in Table 3 confirms the importance of self and mutual contagion coefficients.

The exact calculation of the total log-likelihood would require to estimate 2820 pdf by DFT given that the probability density function of r_t depends on λ_t^1, λ_t^2 and does not admit a closed form expression. As this is computationally too intensive, the total log-likelihood of the model is instead approached by the following expression:

$$L = \sum_{i=1}^{n} \ln\Big[1_{\Delta r_i \geq 0}\left(\lambda_{t_i}^1\,\Delta t + \left(1-\left(\lambda_{t_i}^2\,\Delta t\right)\right)\right)\,\nu_1(\Delta r_i) + \\ +1_{\Delta r_i<0}\left(\lambda_{t_i}^2\,\Delta t + \left(1-\left(\lambda_{t_i}^1\,\Delta t\right)\right)\right)\,\nu_2(\Delta r_i)\Big],$$

and is equal to 23 455 whereas the AIC (Akaike Information criterion) is -46 891. The model is first benchmarked with the model of Vasicek (1977), fitted to the same data set. This model postulates that the dynamics of r_t is driven by the following SDE:

$$dr_t = a(b-r_t)dt + \sigma dW_t \tag{43}$$

where a, b, σ are constant and W_t is a Brownian motion.

The results of the calibration procedure are provided in Table 4. As the one year swap rate has decreased over the period 2004 to 2014 to nearly reach a null value, the level of mean reversion is negative and the speed of mean reversion is close to zero. The log-likelihood for this model is equal to 19 078 and its AIC is -38 150. The comparison of AIC confirms that our model outperforms the Vasicek model (using the BIC leads to the same conclusion). The first graph of Fig. 3 shows the QQ plot of normed residuals for the Vasicek model versus a normal distribution. This plot, as the Jarque-Bera test, confirms the non normality of interest rate variations. [2]

We mention in Section 2, that the model behaves at medium term like a Brownian motion with a random volatility. In this sense, the model is close to GARCH models of Engle (1982). For this reason, we also compare its performance with a GARCH(1,0) model to variations of interest rates. In this framework, the dynamics of short term rates in discrete is the following:

$$\Delta r_t = \mu + \sigma_t \epsilon_t$$

[2] According to Bacry et al. (2013) or Zhu (2013a), the Hawkes process behaves like a Brownian motion for large time scaling. However, our result demonstrates that there is some advantage to model the short term rate by a Hawkes dynamics at the micro-level instead of using a diffusion approximation.

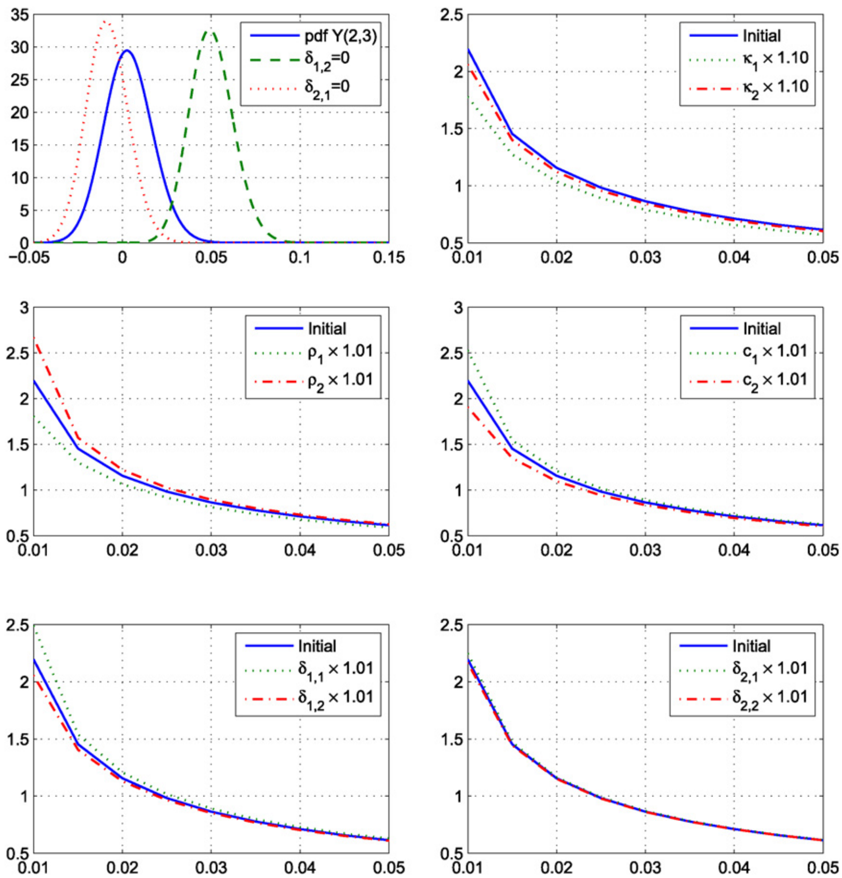

Fig. 8. Sensitivity of curves of implied volatilities to parameters changes, for a set of 2 years caplets, with a 1 year tenor.

where $\epsilon \sim N(0,1)$. The variance σ_t^2 depends both on Δr_{t-1} and on σ_{t-1}^2:

$$\sigma_t^2 = \alpha_0 + \beta_1 \sigma_{t-1}^2.$$

The results of the fit are presented in Table 5. The GARCH(1,0) performs better than the Vasicek model, but its log likelihood is lower than the one of our models. We also test GARCH(2,0) to GARCH (8,0) but the log likelihood never exceeds the one of our models (the maximum obtained is 19 639 with a GARCH(8,0)).

Since the inception of the euro in 1999 and the resulting elimination of exchange-rate risk, swap rates reflect the fluctuations of compensations demanded by AA rated financial institutions for holding mainly liquidity risks. Liquidity risk arises from the potential difficulty to find a counterpart to close a trade relatively quickly.Aït-Sahalia et al. (2015) measure stocks market stress with Hawkes jumps intensities and mention that they reflect market conditions at the time. In a similar way, anticipating poor liquidity conditions is possible through the analysis of filtered intensities, as shown in the second graph of Fig. 4. The frequency of supply orders λ_t^1 is most of the time around the frequency of demand orders λ_t^2, and they have symmetric patterns of evolution. When λ_t^2 is far below λ_t^1, bid orders are not enough frequent compared to ask orders and the market is threatened by a liquidity shortfall. This scenario, happens from the 14/08/2007 to the 09/12/2008, the period that corresponds to the credit crunch crisis.

The econometric calibration is based on historical data and parameters obtained by a such approach define the dynamics of r_t under the real measure of probability, P. Under the risk neutral measure, these parameters depend on some unknown risk premiums θ_1 and θ_2, defining the change of measure (Eq. (24)). On a given date, it is possible to retrieve these premiums by minimizing the sum of spreads between model and observed yield curves. For this purpose, the model yield curve is built with historical parameters of Table 2, adjusted under Q, according to Corollary 3.4. In practice, it is however not relevant to assume that risk premiums are constant over a period of 10 years, as they are directly related to the level of risk aversion in financial markets. For this reason, we recomputed risk premiums at regular intervals of five days of trading. The first subplot of Fig. 5 shows their evolution: θ_1 and θ_2 are respectively positive and negative and nearly symmetric till 2013. The second subplot presents the evolution of $\psi_1(a_1\delta_{1,1} + a_2\delta_{2,1} + \theta_1)$ and $\psi_2(a_2\delta_{2,2} + a_1\delta_{1,2} + \theta_2)$ that multiply parameters c_1, c_2 and $\delta_{i,j}$ under Q, as stated in the Corollary 3.4. This graph reveals that parameters driving λ_t^1 (resp. λ_t^2) are always increased (resp. decreased) under Q. And the steepness of the yield curve is directly related to the distance between $\psi_1(a_1\delta_{1,1} + a_2\delta_{2,1} + \theta_1)$ and $\psi_2(a_2\delta_{2,2} + a_1\delta_{1,2} + \theta_2)$. Around the credit crunch, yield curves are indeed nearly flat and $\psi_1(a_1\delta_{1,1} + a_2\delta_{2,1} + \theta_1)$ and $\psi_2(a_2\delta_{2,2} + a_1\delta_{1,2} + \theta_2)$ are very close to one.

Table 6 shows the parameters under Q on the 28/11/2014. The first subplot of Fig. 6 compares the yields produced by the model for these parameters (line labeled "Model Q'") with the observed ones. If we replace κ_1, κ_2, c_1^Q and c_2^Q by values in the column "Best fit" of Table 6, the model replicates perfectly the market data. This is illustrated in the first subplot of Fig. 6 by the curve labeled "Best fit". These parameters cannot be reconciled anymore with historical parameters, through an affine change of measure but does not appear irrelevant. We use them later to analyze the sensitivity of the model to each of its parameters.

The five last subplots of the Fig. 6 shows the marginal effect of each parameter on the slope of the yield curve produced by the model. Increasing c_1, $\delta_{1,1}$ or $\delta_{1,2}$ raises the frequency of positive variations of the interest rate, and then the steepness of the curve. Increasing c_2, $\delta_{2,2}$ or $\delta_{2,1}$ steps up the frequency of negative variations of the short term rate and flatten the yield curve. Using higher speeds of mean reversion, $\kappa_1$and κ_2, have the same effect on the curve. As the average size of orders are inversely proportional to ρ_1 and ρ_2, increasing these parameters is equivalent to decrease the average amplitude of variations of the interest rate. Then higher ρ_1 or ρ_2 respectively lowers or raises the steepness of the curve.

Fig. 7 presents simulated sample paths for r_t, λ_t^1 and λ_t^2 and their mean calculated under Q, with Propositions 2.1 and 2.2. Simulated paths depict periods of decline, sharp increase and stability, that are comparable to real ones shown in Fig. 4. As mentioned earlier, the model is not explicitly mean reverting. However the simulation of r_t reveals divergences to extreme value is avoided (at least at medium term) by the mutual dependence between arrivals of bid and ask orders.

We also observe negative short term rates in two scenarios during the first five years, as the expected short term rate is close to 0%. In fact, the number of scenarios in which negative rates are generated, directly depends on parameters of mutual excitations $\delta_{1,2}$ and $\delta_{2,1}$. The higher is $\delta_{1,2}$, the higher is the probability of observing an upward jump following a downward variation of interest rates and the lower is the probability of observing negative short term rates. This point is emphasized by the first subplot of Fig. 8, that presents the probability density function (computed by DFT) of the forward yield, $Y(2,3)$ as defined by Eq. (38), for different levels of cross excitations. We see that setting $\delta_{1,2}$ to zero is enough to exclude negative forward yields.

The five last subplots of Fig. 8 show selected curves of implied volatilities, for a set of 2 years caplets, with a 1 year tenor. Prices are obtained by a Fourier transform with $M = 2^{12}$ steps of discretization and $y_{max} = 0.10$. Implied volatilities are next obtained by inverting the Black & Scholes formula for caplets. The purpose of these graphs is to illustrate the sensitivity of implied volatilities to a change of key parameters defining the model. Parameters are these fitting market on the 28/11/2014. Increasing c_1, $\delta_{1,1}$ or $\delta_{1,2}$ increases the steepness of the smile of volatilities. Whereas increasing c_2, $\delta_{2,2}$ or $\delta_{2,1}$ flatten the curve. Finally, higher ρ_1 or ρ_2 respectively lowers or raises the slope of the smile.

5. Conclusion

The literature provides a great deal of evidence that liquidity shortages are caused by a disequilibrium between the supply and demand of fixed income instruments, and that it impacts level of interest rates. Directly inspired from the economic monetary theory, this work presents a new interest rate model based on recent developments in the study of market micro-structure. The novelty of this approach is to consider that aggregate supply and demand of fixed income products are ruled by a bivariate Hawkes process. This introduces both path dependence and mutual excitation in the arrivals processes of bid and ask orders for interest rate products. Furthermore, quasi explicit formulas are available for moments of intensities and bond prices and changes of measure.

The econometric analysis of the one year swap rate over a period of 10 years, suggests that intensities of bid/ask orders arrivals are key factors to understand the fluctuations of rates. In particular, a negative difference between bid and ask frequencies is a solid indicator to detect liquidity shortfalls. On the other hand, combining the econometric calibration with the analysis of past swap curves, allows us to filter risk premiums of processes representing the demand and supply of bonds. The distance between these risk premiums explains the steepness of the yield curve and is particularly small during the 2008 crisis. Finally, the different sensitivity analysis developed in this work, confirm that the model is tractable for derivatives pricing or for risk management purposes.

6. Acknowledgment

I thank the two anonymous referees and the editor for their comments. They clearly contributed to improve the paper. I also thank the Chair of “Data Analytics and Models for Insurance” of BNP Paribas Cardiff, hosted by ISFA (Université Claude Bernard, Lyon).

Appendix A. Proofs of propositions

A.1. Proposition 2.1

Proof. Consider the functions $g^i = \lambda_t^i$ for $i = 1, 2$. According to Eqs. (11) and (12), their expectations are such that

$$\begin{aligned}
\mathbb{E}\left(\mathcal{A}g^1\right) &= \kappa_1\left(c_1 - \mathbb{E}\left(\lambda_t^1\right)\right) + \mathbb{E}\left(\lambda_t^1\right)\int_{-\infty}^{+\infty}\delta_{1,1}z\,d\nu_1(z) + \mathbb{E}\left(\lambda_t^2\right) \\
&\quad\times\int_{-\infty}^{+\infty}\delta_{1,2}z\,d\nu_2(z) \\
&= \kappa_1\left(c_1 - \mathbb{E}\left(\lambda_t^1\right)\right) + \mathbb{E}\left(\lambda_t^1\right)\delta_{1,1}\mu_1 + \mathbb{E}\left(\lambda_t^2\right)\delta_{1,2}\mu_2 \\
\mathbb{E}\left(\mathcal{A}g^2\right) &= \kappa_2\left(c_2 - \mathbb{E}\left(\lambda_t^2\right)\right) + \mathbb{E}\left(\lambda_t^1\right)\int_{-\infty}^{+\infty}\delta_{2,1}z\,d\nu_1(z) + \mathbb{E}\left(\lambda_t^2\right) \\
&\quad\times\int_{-\infty}^{+\infty}\delta_{2,2}z\,d\nu_2(z) \\
&= \kappa_2\left(c_2 - \mathbb{E}\left(\lambda_t^2\right)\right) + \mathbb{E}\left(\lambda_t^1\right)\delta_{2,1}\mu_1 + \mathbb{E}\left(\lambda_t^2\right)\delta_{2,2}\mu_2
\end{aligned}$$

If we refer to Eq. (13), moments $m_1(t)$ and $m_2(t)$ are solutions of a system of ordinary differential equations (ODEs) with respect to time.

$$\frac{\partial}{\partial t}\begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \begin{pmatrix} \kappa_1 c_1 \\ \kappa_2 c_2 \end{pmatrix} + \begin{pmatrix} (\delta_{1,1}\mu_1 - \kappa_1) & \delta_{1,2}\mu_2 \\ \delta_{2,1}\mu_1 & (\delta_{2,2}\mu_2 - \kappa_2) \end{pmatrix}\begin{pmatrix} m_1 \\ m_2 \end{pmatrix} \tag{44}$$

Finding a solution requires to determine eigenvalues γ and eigenvectors (v_1, v_2) of the matrix present in the right term of this system.

$$\begin{pmatrix} (\delta_{1,1}\mu_1 - \kappa_1) & \delta_{1,2}\mu_2 \\ \delta_{2,1}\mu_1 & (\delta_{2,2}\mu_2 - \kappa_2) \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \gamma\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

Eigenvalues cancel the determinant of the following matrix:

$$\det\begin{pmatrix} (\delta_{1,1}\mu_1 - \kappa_1) - \gamma & \delta_{1,2}\mu_2 \\ \delta_{2,1}\mu_1 & (\delta_{2,2}\mu_2 - \kappa_2) - \gamma \end{pmatrix} = 0$$

and are solutions of the following second order equation:

$$\begin{aligned}
&\gamma^2 - \gamma\left((\delta_{1,1}\mu_1 - \kappa_1) + (\delta_{2,2}\mu_2 - \kappa_2)\right) + (\delta_{1,1}\mu_1 - \kappa_1)(\delta_{2,2}\mu_2 - \kappa_2) \\
&\quad - \delta_{1,2}\delta_{2,1}\mu_1\mu_2 = 0.
\end{aligned}$$

Roots of this last equation are γ_1 and γ_2, as defined by Eq. (15). One way to find an eigenvector is to note that it must be orthogonal to each rows of the following matrix:

$$\begin{pmatrix} (\delta_{1,1}\mu_1-\kappa_1)-\gamma & \delta_{1,2}\mu_2 \\ \delta_{2,1}\mu_1 & (\delta_{2,2}\mu_2-\kappa_2)-\gamma \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}=0,$$

then necessarily,

$$\begin{pmatrix} v_1^i \\ v_2^i \end{pmatrix}=\begin{pmatrix} -\delta_{1,2}\mu_2 \\ (\delta_{1,1}\mu_1-\kappa_1)-\gamma_i \end{pmatrix} \quad \text{for } i=1,2.$$

Let $D = diag(\gamma_1,\gamma_2)$. The matrix in the right term of Eq. (44) admits the representation

$$\begin{pmatrix} (\delta_{1,1}\mu_1-\kappa_1) & \delta_{1,2}\mu_2 \\ \delta_{2,1}\mu_1 & (\delta_{2,2}\mu_2-\kappa_2) \end{pmatrix}=VDV^{-1},$$

where V is the matrix of eigenvectors, as defined in Eq. (16). Its determinant, Υ, and its inverse are respectively provided by Eqs. (18) and (17). Two new variables are defined as follows:

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}=V^{-1}\begin{pmatrix} m_1 \\ m_2 \end{pmatrix}.$$

The system (Eq. (44)) is decoupled into two independent ODEs:

$$\frac{\partial}{\partial t}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}=V^{-1}\begin{pmatrix} \kappa_1 c_1 \\ \kappa_2 c_2 \end{pmatrix}+\begin{pmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \tag{45}$$

And introducing the following notations:

$$V^{-1}\begin{pmatrix} \kappa_1 c_1 \\ \kappa_2 c_2 \end{pmatrix}=\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix},$$

leads to the solutions for the system (Eq. (45)):

$$\begin{aligned} u_1(t) &= \frac{\epsilon_1}{\gamma_1}\left(e^{\gamma_1 t}-1\right)+d_1 e^{\gamma_1 t} \\ u_2(t) &= \frac{\epsilon_2}{\gamma_2}\left(e^{\gamma_2 t}-1\right)+d_2 e^{\gamma_2 t} \end{aligned}$$

where $d=(d_1,d_2)'$ is such that $d=V^{-1}\lambda_0$. Or in matrix form,

$$\begin{aligned} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} =& \begin{pmatrix} \frac{1}{\gamma_1}\left(e^{\gamma_1 t}-1\right) & 0 \\ 0 & \frac{1}{\gamma_2}\left(e^{\gamma_2 t}-1\right) \end{pmatrix} V^{-1}\begin{pmatrix} \kappa_1 c_1 \\ \kappa_2 c_2 \end{pmatrix} \\ &+\begin{pmatrix} e^{\gamma_1 t} & 0 \\ 0 & e^{\gamma_2 t} \end{pmatrix} V^{-1}\begin{pmatrix} \lambda_0^1 \\ \lambda_0^2 \end{pmatrix}. \end{aligned}$$

Expressions (Eq. (14)) for m_1, m_2 are inferred from this last relation. □

A.2. Corollary 2.2

Proof. Let us denote $f=\mathbb{E}(X_t)$, then the expectation of its infinitesimal generator is equal to

$$\begin{aligned} \mathbb{E}(\mathcal{A}f) &= \mathbb{E}\left(\lambda_t^1\right)\int_{-\infty}^{+\infty}\alpha_1 z\, d\nu_1(z)-\mathbb{E}\left(\lambda_t^2\right)\int_{-\infty}^{+\infty}\alpha_2 z\, d\nu_2(z) \\ &= \mathbb{E}\left(\lambda_t^1\right)\alpha_1\mu_1-\mathbb{E}\left(\lambda_t^2\right)\alpha_2\mu_2 \end{aligned}$$

and according to Eq. (12), we conclude. □

A.3. Proposition 2.3

Proof. Let us introduce the following notations: $g_i=\left(\lambda_t^i\right)^2$ for $i=1,2$ and $g_3=\lambda_t^1\lambda_t^2$, according to Eqs. (18) and (17), the next relation holds

$$\begin{aligned} \mathcal{A}g_1 =& 2\lambda_t^1\kappa_1 c_1-2\left(\lambda_t^1\right)^2\kappa_1+\lambda_t^1\int_{-\infty}^{+\infty}2\lambda_t^1\delta_{1,1}z+(\delta_{1,1}z)^2\, d\nu_1(z) \\ &+\lambda_t^2\int_{-\infty}^{+\infty}2\lambda_t^1\delta_{1,2}z+(\delta_{1,2}z)^2\, d\nu_2(z) \\ \mathcal{A}g_2 =& 2\lambda_t^2\kappa_2 c_2-2\left(\lambda_t^2\right)^2\kappa_2+\lambda_t^2\int_{-\infty}^{+\infty}2\lambda_t^2\delta_{2,2}z+(\delta_{2,2}z)^2\, d\nu_2(z) \\ &+\lambda_t^1\int_{-\infty}^{+\infty}2\lambda_t^2\delta_{2,1}z+(\delta_{2,1}z)^2\, d\nu_1(z) \\ \mathcal{A}g_3 =& \kappa_1(c_1-\lambda_t^1)\lambda_t^2+\kappa_2(c_2-\lambda_t^2)\lambda_t^1 \\ &+\lambda_t^1\int_{-\infty}^{+\infty}\left(\lambda_t^1+\delta_{1,1}z\right)\left(\lambda_t^2+\delta_{2,1}z\right)-\lambda_t^1\lambda_t^2\, d\nu_1(z) \\ &+\lambda_t^2\int_{-\infty}^{+\infty}\left(\lambda_t^1+\delta_{1,2}z\right)\left(\lambda_t^2+\delta_{2,2}z\right)-\lambda_t^1\lambda_t^2\, d\nu_2(z). \end{aligned}$$

If we note $v_i=\mathbb{E}\left(\left(\lambda_t^i\right)^2\right)$ for $i=1,2$ and $v_3=\mathbb{E}\left(\lambda_t^1\lambda_t^2\right)$, a system of ODEs is deduced from Eq. (13):

$$\begin{aligned} \frac{\partial}{\partial t}v_1 =& 2m_1(t)\kappa_1 c_1-2v_1(t)\kappa_1+2v_2(t)\delta_{1,1}\mu_1 \\ &+m_1(t)\delta_{1,1}^2\eta_1+2v_3(t)\delta_{1,2}\mu_2+m_2(t)\delta_{1,2}^2\eta_2 \\ \frac{\partial}{\partial t}v_2 =& 2m_2(t)\kappa_2 c_2-2v_2(t)\kappa_2+2v_2(t)\delta_{2,2}\mu_2 \\ &+m_2(t)\delta_{2,2}^2\eta_2+2v_3(t)\delta_{2,1}\mu_1+m_1(t)\delta_{2,1}^2\eta_1 \\ \frac{\partial}{\partial t}v_3 =& m_2(t)\kappa_1 c_1-\kappa_1 v_3(t)+m_1(t)\kappa_2 c_2-\kappa_2 v_3(t) \\ &+v_1(t)\delta_{2,1}\mu_1+v_3(t)\delta_{1,1}\mu_1+m_1(t)\delta_{1,1}\delta_{2,1}\eta_1 \\ &+v_2(t)\delta_{1,2}\mu_2+v_3(t)\delta_{2,2}\mu_2+m_2(t)\delta_{2,2}\delta_{1,2}\eta_2. \end{aligned} \tag{46}$$

As centered second moments $V_i(t)$, are linked to non centered ones, v_i by the next differential equations:

$$\begin{cases} \frac{\partial}{\partial t}V_i=\frac{\partial}{\partial t}v_i-2m_i\frac{\partial}{\partial t}m_i \quad i=1,2 \\ \frac{\partial}{\partial t}V_3=\frac{\partial}{\partial t}v_3-m_1\frac{\partial}{\partial t}m_2-m_2\frac{\partial}{\partial t}m_1. \end{cases}$$

It is sufficient to combine Eqs. (44) and (46) to conclude. □

A.4. Proposition 2.4

Proof. Let us define $f\left(t,\lambda_t^1,J_t^1,\lambda_t^2,J_t^2\right) := \mathbb{E}\left(e^{w_0 X_T - w_1\int_t^T X_s ds + \begin{pmatrix} w_2 \\ w_3\end{pmatrix}^\top \begin{pmatrix}\lambda_T^1 \\ \lambda_T^2\end{pmatrix}} \mid \mathcal{F}_t\right)$ and let f_t, f_{λ^1} and f_{λ^2} denote respectively the partial derivatives of f with respect to time and intensities. According to the Feynman-Kac formula, $f(.)$ is solution of the next partial integro differential equation:

$$\begin{aligned}
w_1\left(\alpha_1 L_t^1-\alpha_2 L_t^2\right)f &= f_t+\kappa_1\left(c_1-\lambda_t^1\right)f_{\lambda^1}+\kappa_2\left(c_2-\lambda_t^2\right)f_{\lambda^2} \\
&+\lambda_t^1\int_{-\infty}^{+\infty} f\left(t,\lambda_t^1+\delta_{1,1}z,J_t^1+(z,1)^\top,\lambda_t^2+\delta_{2,1}z,J_t^2\right) \\
&-f\left(t,\lambda_t^1,J_t^1,\lambda_t^2,J_t^2\right)d\nu_1(z) \\
&+\lambda_t^2\int_{-\infty}^{+\infty} f\left(t,\lambda_t^1+\delta_{1,2}z,J_t^1,\lambda_t^2+\delta_{2,2}z,J_t^2+(z,1)^\top\right) \\
&-f\left(t,\lambda_t^1,J_t^1,\lambda_t^2,J_t^2\right)d\nu_2(z). \qquad (47)
\end{aligned}$$

f also satisfies the next limit condition

$$f\left(T,\lambda_T^1,J_T^1,\lambda_T^2,J_T^2\right)=\exp\left(w_0\left(\alpha_1 L_T^1-\alpha_2 L_T^2\right)+\begin{pmatrix} w_2 \\ w_3\end{pmatrix}^\top\begin{pmatrix}\lambda_T^1 \\ \lambda_T^2\end{pmatrix}\right) \qquad (48)$$

Let us assume that f has an exponential form

$$f=\exp\left(A(t,T)+B(t,T)^\top\begin{pmatrix}\lambda_t^1 \\ \lambda_t^2\end{pmatrix}+C(t,T)^\top\begin{pmatrix}L_t^1 \\ L_t^2\end{pmatrix}\right)$$

where $B(t,T)^\top=(B_1(t,T),B_2(t,T))^\top$ and $C(t,T)^\top=(C_1(t,T),C_2(t,T))^\top$. Under this assumption, Eq. (47) becomes

$$\begin{aligned}
w_1\left(\alpha_1 L_t^1-\alpha_2 L_t^2\right) &= \left(\frac{\partial}{\partial t}A+\lambda_t^1\frac{\partial}{\partial t}B_1+\lambda_t^2\frac{\partial}{\partial t}B_2+L_t^1\frac{\partial}{\partial t}C_1+L_t^2\frac{\partial}{\partial t}C_2\right) \\
&+\lambda_t^1\left[\left(\left(\psi_1(B_1\delta_{1,1}+C_1+B_2\delta_{2,1})\right)-1\right]+\kappa_1\left(c_1-\lambda_t^1\right)B_1 \\
&+\lambda_t^2\left[\left(\psi_2\left(B_1\delta_{1,2}+C_2+B_2\delta_{2,2}\right)-1\right)\right]+\kappa_2\left(c_2-\lambda_t^2\right)B_2. \qquad (49)
\end{aligned}$$

Since this relation holds for any λ_t^i and $L_t^i i$=1,2, this last relation holds only if their multiplicative coefficients are null. This is achieved only if

$$\begin{aligned}
0 &= \frac{\partial}{\partial t}B_1-\kappa_1 B_1+\left[\psi_1\left(B_1\delta_{1,1}+C_1+B_2\delta_{2,1}\right)-1\right], \\
0 &= \frac{\partial}{\partial t}B_2-\kappa_2 B_2+\left[\psi_2\left(B_1\delta_{1,2}+C_2+B_2\delta_{2,2}\right)-1\right], \\
0 &= \frac{\partial}{\partial t}A+\kappa_1 c_1 B_1+\kappa_2 c_2 B_2, \\
w_1\alpha_1 &= \frac{\partial}{\partial t}C_1-w_1\alpha_2=\frac{\partial}{\partial t}C. \qquad (50)
\end{aligned}$$

From the boundary condition (Eq. (48)), we infer that $C_1(t,T)=w_0\alpha_1-w_1\alpha_1(T-t)$ and $C_2(t,T)=-w_0\alpha_2+w_1\alpha_2(T-t)$. □

A.5. Proposition 3.1

Proof. Let us denote by Y_t the exponent of M_t:

$$Y_t=(a_1(\theta_1,\theta_2),\ a_2(\theta_1,\theta_2))\begin{pmatrix}\lambda_t^1 \\ \lambda_t^2\end{pmatrix}+(\theta_1,\theta_2)\begin{pmatrix}L_t^1 \\ L_t^2\end{pmatrix}-\varphi(\theta_1,\theta_2)t. \qquad (51)$$

According to Eq. (10), its infinitesimal dynamics is given by

$$\begin{aligned}
dY_t &= a_1\kappa_1\left(c_1-\lambda_t^1\right)dt+a_2\kappa_2\left(c_2-\lambda_t^2\right)dt+\left(a_1\delta_{1,1}+a_2\delta_{2,1}+\theta_1\right)dL_t^1 \\
&+\left(a_2\delta_{2,2}+a_1\delta_{1,2}+\theta_2\right)dL_t^2-\varphi(\theta_1,\theta_2)dt.
\end{aligned}$$

In the remainder of this proof, the random measure of O^i is noted $\chi^i(.)$ and is such that $O^i=\int_{-\infty}^{\infty}\chi^i(dz)$ for $i=1,2$. Applying the Ito's lemma for semi-martingales to M_t leads to the next relation:

$$\begin{aligned}
dM_t =& M_t dY_t+\frac{1}{2}M_t d\left[Y_t,Y_t\right]_t^c \\
&+M_t\int_{-\infty}^{\infty}\left(e^{\left(a_1\delta_{1,1}+a_2\delta_{2,1}+\theta_1\right)z}-1-\left(a_1\delta_{1,1}+a_2\delta_{2,1}+\theta_1\right)z\right)\chi^1(dz)dN_t^1 \\
&+M_t\int_{-\infty}^{\infty}\left(e^{\left(a_2\delta_{2,2}+a_1\delta_{1,2}+\theta_2\right)z}-1-\left(a_2\delta_{2,2}+a_1\delta_{1,2}+\theta_2\right)z\right)\chi^2(dz)dN_t^2
\end{aligned}$$

or equal to

$$\begin{aligned}
dM_t =& M_t\left(a_1\kappa_1 c_1+a_2\kappa_2 c_2-\varphi\right)dt \\
&-M_t\lambda_t^1\left(a_1\kappa_1-\int_{-\infty}^{\infty}\left(e^{\left(a_1\delta_{1,1}+a_2\delta_{2,1}+\theta_1\right)z}-1\right)\nu_1(dz)\right)dt \\
&-M_t\lambda_t^2\left(a_2\kappa_2-\int_{-\infty}^{\infty}\left(e^{\left(a_2\delta_{2,2}+a_1\delta_{1,2}+\theta_2\right)z}-1\right)\nu_2(dz)\right)dt \\
&+M_t\int_{-\infty}^{\infty}\left(e^{\left(a_1\delta_{1,1}+a_2\delta_{2,1}+\theta_1\right)z}-1\right)\left[\chi^1(dz)dN_t^1-\lambda_t^1\nu_1(dz)dt\right] \\
&+M_t\int_{-\infty}^{\infty}\left(e^{\left(a_2\delta_{2,2}+a_1\delta_{1,2}+\theta_2\right)z}-1\right)\left[\chi^2(dz)dN_t^2-\lambda_t^2\nu_2(dz)dt\right].
\end{aligned}$$

Since the integrals with respect to $\chi^i(dz)dN_t^i-\lambda_t^i\nu_i(dz)dt$ are local martingales, M_t is also a local martingale if and only if the following relations hold:

$$\begin{cases}
a_1(\theta_1,\theta_2)\kappa_1 c_1+a_2(\theta_1,\theta_2)\kappa_2 c_2-\varphi(\theta_1,\theta_2) & =0 \\
a_1(\theta_1,\theta_2)\kappa_1-\int_{-\infty}^{\infty}\left(e^{\left(a_1(\theta_1,\theta_2)\delta_{1,1}+a_2(\theta_1,\theta_2)\delta_{2,1}+\theta_1\right)z}-1\right)\nu_1(dz) & =0 \\
a_2(\theta_1,\theta_2)\kappa_2-\int_{-\infty}^{\infty}\left(e^{\left(a_2(\theta_1,\theta_2)\delta_{2,2}+a_1(\theta_1,\theta_2)\delta_{1,2}+\theta_2\right)z}-1\right)\nu_2(dz) & =0
\end{cases}$$

and these conditions are equivalent Eqs. (25) and (26). □

A.6. Proposition 3.2

Proof. If Y_t is the exponent of M_t, as defined by Eq. (51), the moment generating function of X_T under the risk neutral is then equal to

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\left(e^{wX_T}|\mathcal{F}_t\right) &= \mathbb{E}\left(e^{Y_T-Y_t+wX_T}|\mathcal{F}_t\right) \\
&= e^{-Y_t}\mathbb{E}\left(e^{Y_T+wX_T}|\mathcal{F}_t.\right)
\end{aligned}$$

If $f\left(t,\lambda_t^1,J_r^1,\lambda_r^2,J_r^2\right)$ denotes $\mathbb{E}\left(e^{Y_T+wX_T}|\mathcal{F}_t\right)$, according to the Itô's lemma, it solves the next equation

$$\begin{aligned}
0 =& f_t+\kappa_1\left(c_1-\lambda_t^1\right)f_{\lambda^1}+\kappa_2\left(c_2-\lambda_t^2\right)f_{\lambda^2} \\
&+\lambda_t^1\int_{-\infty}^{+\infty} f\left(t,\lambda_t^1+\delta_{1,1}z,J_t^1+(z,1)^\top,\lambda_t^2+\delta_{2,1}z,J_t^2\right) \\
&-f\left(t,\lambda_t^1,J_t^1,\lambda_t^2,J_t^2\right)d\nu_1(z) \\
&+\lambda_t^2\int_{-\infty}^{+\infty} f\left(t,\lambda_t^1+\delta_{1,2}z,J_t^1,\lambda_t^2+\delta_{2,2}z,J_t^2+(z,1)^\top\right) \\
&-f\left(t,\lambda_t^1,J_t^1,\lambda_t^2,J_t^2\right)d\nu_2(z). \qquad (52)
\end{aligned}$$

where $f_t, f_{\lambda^1}, f_{\lambda^2}$ are the partial derivatives of $f(.)$ with respect to time and intensities. Furthermore given that

$$\begin{aligned}Y_T + wX_T =& a_1\left(\lambda_T^1 - \kappa_1 c_1 T\right) + a_2\left(\lambda_T^2 - \kappa_2 c_2 T\right)\\ &+ (\theta_1 + \alpha_1 w)\, L_T^1 + (\theta_2 + \alpha_2 w)\, L_T^2\,,\end{aligned} \tag{53}$$

$f(.)$ satisfies the following terminal condition at time $t = T$:

$$\begin{aligned}f\left(T, \lambda_T^1, J_T^1, \lambda_T^2, J_T^2\right) =& \exp\Big((\theta_1 + \alpha_1 w)\, L_T^1 + (\theta_2 + \alpha_2 w)\, L_T^2\\ &+ a_1\left(\lambda_T^1 - \kappa_1 c_1 T\right) + a_2\left(\lambda_T^2 - \kappa_2 c_2 T\right)\Big)\,.\end{aligned}$$

In the remainder of this section, it is assumed that $f(.)$ is an exponential affine function:

$$f = \exp\left(A(t,T) + B(t,T)^\top \begin{pmatrix}\psi_1(a_1\delta_{1,1} + a_2\delta_{2,1} + \theta_1)\lambda_t^1\\ \psi_2(a_2\delta_{2,2} + a_1\delta_{1,2} + \theta_2)\lambda_t^2\end{pmatrix} + C(t,T)^\top \begin{pmatrix}L_t^1\\ L_t^2\end{pmatrix}\right)$$

where $B(t,T,w)^\top = (B_1(t,T,w), B_2(t,T,w))^\top$ and $C(t,T,w)^\top = (C_1(t,T,w), C_2(t,T,w))^\top$. Under this assumption, the partial derivatives of f are given by

$$f_t = \left(\frac{\partial}{\partial t}A + \psi_1\lambda_t^1\frac{\partial}{\partial t}B_1 + \psi_2\lambda_t^2\frac{\partial}{\partial t}B_2 + L_t^1\frac{\partial}{\partial t}C_1 + L_t^2\frac{\partial}{\partial t}C_2\right)f$$

$$f_{\lambda^1} = \psi_1 B_1 f \quad f_{\lambda^2} = \psi_2 B_2 f$$

where ψ_1 and ψ_2 abusively denote $\psi_1(a_1\delta_{1,1} + a_2\delta_{2,1} + \theta_1)$ and $\psi_2(a_2\delta_{2,2} + a_1\delta_{1,2} + \theta_2)$. Integrands in Eq. (52) are equal to

$$\begin{aligned}&f\left(t, \lambda_t^1 + \delta_{1,1}z, J_t^1 + (z,1)^\top, \lambda_t^2 + \delta_{2,1}z, J_t^2\right) - f\left(\lambda_t^1, J_t^1, \lambda_t^2, J_t^2\right)\\ &\quad = f\left[\exp\left((B_1\psi_1\delta_{1,1} + C_1 + B_2\psi_2\delta_{2,1})\, z\right) - 1\right],\\ &f\left(\lambda_t^1 + \delta_{1,2}z, J_t^1, \lambda_t^2 + \delta_{2,2}z, J_t^2 + (z,1)^\top\right) - f\left(\lambda_t^1, J_t^1, \lambda_t^2, J_t^2\right)\\ &\quad = f\left[\exp\left((B_1\psi_1\delta_{1,2} + C_2 + B_2\psi_2\delta_{2,2})\, z\right) - 1\right].\end{aligned}$$

Injecting these expressions in Eq. (52) yields a system of ODEs for A, B_1 and B_2:

$$\begin{aligned}0 &= \frac{\partial}{\partial t}B_1 - \kappa_1 B_1 + \frac{1}{\psi_1}\left[\psi_1\left(B_1\psi_1\delta_{1,1} + C_1 + B_2\psi_2\delta_{2,1}\right) - 1\right],\\ 0 &= \frac{\partial}{\partial t}B_2 - \kappa_2 B_2 + \frac{1}{\psi_2}\left[\psi_2\left(B_1\psi_1\delta_{1,2} + C_2 + B_2\psi_2\delta_{2,2}\right) - 1\right],\\ 0 &= \frac{\partial}{\partial t}A + \psi_1\kappa_1 c_1 B_1 + \psi_2\kappa_2 c_2 B_2,\\ 0 &= \frac{\partial}{\partial t}C_1, \quad 0 = \frac{\partial}{\partial t}C_2\end{aligned} \tag{54}$$

with the terminal conditions:

$$\begin{cases}A(T,T) &= -a_1\kappa_1 c_1 T - a_2\kappa_2 c_2 T\\ B_1(T,T) &= \frac{a_1}{\psi_1} \quad B_2(T,T) = \frac{a_2}{\psi_2}\\ C_1(T,T) &= (\theta_1 + \alpha_1 w)\\ C_2(T,T) &= (\theta_2 + \alpha_2 w)\,.\end{cases}$$

As $C_1(t,T) = \theta_1 + \alpha_1 w$ and $C_2(t,T) = \theta_2 + \alpha_2 w$, the moment generating function of X_t is equal to

$$\begin{aligned}\mathbb{E}^Q\left(e^{wX_T}|\mathcal{F}_t\right) &= e^{-Y_t}\mathbb{E}\left(e^{Y_T + wX_T}|\mathcal{F}_t\right)\\ &= \exp\Big(A + a_1\kappa_1 c_1 t + a_2\kappa_2 c_2 t + (B_1\psi_1 - a_1)\,\lambda_t^1\\ &\quad + (B_2\psi_2 - a_2)\,\lambda_t^2 + wX_t\Big)\,.\end{aligned}$$

In the remainder of the proof, this expectation is restated in a form similar to the moment generating function of X_T under P. To achieve this, the following change of variables is done:

$$\begin{aligned}A' &:= A + a_1\kappa_1 c_1 t + a_2\kappa_2 c_2 t,\\ B_1' &:= B_1 - \frac{a_1}{\psi_1},\\ B_2' &:= B_2 - \frac{a_2}{\psi_2}\end{aligned}$$

with the terminal conditions $A'(T,T) = 0$, $B_1'(T,T) = 0$, $B_2'(T,T) = 0$. As from Eq. (25), the following relation holds:

$$\begin{cases}\kappa_1\frac{a_1}{\psi_1} = \left(1 - \frac{1}{\psi_1}\right)\\ \kappa_2\frac{a_2}{\psi_2} = \left(1 - \frac{1}{\psi_2}\right)\end{cases} \tag{55}$$

The system of ODEs (Eq. (54)) becomes

$$\begin{aligned}0 &= \frac{\partial}{\partial t}B_1' - \kappa_1 B_1' + \Big[\frac{1}{\psi_1}\psi_1\left(B_1'\psi_1\delta_{1,1} + (\theta_1 + \delta_{1,1}a_1 + \delta_{2,1}a_2)\right.\\ &\qquad\left. + \alpha_1 w + B_2'\psi_2\delta_{2,1}\right) - 1\Big],\\ 0 &= \frac{\partial}{\partial t}B_2' - \kappa_2 B_2' + \Big[\frac{1}{\psi_2}\psi_2\left(B_1'\psi_1\delta_{1,2} + (\theta_2 + \delta_{1,2}a_1 + \delta_{2,2}a_2)\right.\\ &\qquad\left. + \alpha_2 w + B_2'\psi_2\delta_{2,2}\right) - 1\Big],\\ 0 &= \frac{\partial}{\partial t}A' + \psi_1\kappa_1 c_1 B_1' + \psi_2\kappa_2 c_2 B_2'.\end{aligned}$$

If we consider jumps $O^{1,Q}$, $O^{2,Q}$ that have moment generating functions defined by Eq. (29),the moment generating function of X_T under Q is given by

$$\mathbb{E}^Q\left(e^{wX_T}|\mathcal{F}_t\right) = \exp\left(wX_t + A'(t,T) + \begin{pmatrix}B_1'(t,T)\\ B_2'(t,T)\end{pmatrix}^\top \begin{pmatrix}\lambda_t^{1,Q}\\ \lambda_t^{2,Q}\end{pmatrix}\right) \tag{56}$$

where A', B_1' and B_2' solve a system, identical to the one of Proposition 2.4. □

A.7. Corollary 3.3

Proof. If we denote $\beta_1 = \delta_{1,1}a_1 + \delta_{2,1}a_2 + \theta_1$, by construction the moment generating function of sell orders, under the risk neutral measure is provided by the following ratio:

$$\begin{aligned}\psi_1^Q(z) &= \frac{\psi_1(z + \beta_1)}{\psi_1(\beta_1)}\\ &= \frac{\rho_1}{\rho_1 - z - \beta_1}\frac{\rho_1 - \beta_1}{\rho_1}\end{aligned}$$

and we conclude that sell orders are also exponential under Q. The same reasoning holds for ask orders. □

A.8. Corollary 3.6

Proof. According to the Ito's lemma for semi-martingales, $P\left(t,T,\lambda_t^{1Q},J_t^{1Q},\lambda_t^{2Q},J_t^{2Q}\right)$ is such that

$$\begin{aligned}dP &= P_t + \kappa_1\left(c_1^Q-\lambda_t^{1Q}\right)P_{\lambda^1}dt + \kappa_2\left(c_2^Q-\lambda_t^{2Q}\right)P_{\lambda^2}dt\\ &\quad+\int_{-\infty}^{+\infty} P\left(t,\lambda_t^{1Q}+\delta_{1,1}^Q z, J_t^{1Q}+(z,1)^\top,\lambda_t^{2Q}+\delta_{2,1}^Q z, J_t^{2Q}\right)\\ &\quad- P\left(t,\lambda_t^{1Q},J_t^{1Q},\lambda_t^{2Q},J_t^{2Q}\right)\, L^{1Q}(dt,dz)\\ &\quad+\int_{-\infty}^{+\infty} P\left(t,\lambda_t^{1Q}+\delta_{1,2}^Q z, J_t^{1Q},\lambda_t^{2Q}+\delta_{2,2}^Q z, J_t^{2Q}+(z,1)^\top\right)\\ &\quad- P\left(t,\lambda_t^{1Q},J_t^{1Q},\lambda_t^{2Q},J_t^{2Q}\right)\, L^{2Q}(dt,dz),\end{aligned} \tag{57}$$

where partial derivatives are obtained from Eqs. (35) and (36) □

A.9. Corollary 3.7

Proof. By definition of the forward measure and using the fact that $\mathcal{F}_t \subset \mathcal{F}_T$, the Laplace transform of $Y(T,S)$ is given by:

$$\begin{aligned}\mathbb{E}^S\left(e^{wY(T,S)}\mid\mathcal{F}_t\right) &= \frac{\mathbb{E}^Q\left(\left(e^{\int_0^S r_s ds}\mathbb{E}^Q\left(e^{-\int_0^S r_s ds}|\mathcal{F}_0\right)\right)^{-1} e^{wY(T,S)}\mid\mathcal{F}_t\right)}{e^{-\int_0^t r_s ds}\mathbb{E}^Q\left(e^{-\int_t^S r_s ds}\,|\mathcal{F}_t\right)\left(\mathbb{E}^Q\left(e^{-\int_0^S r_s ds}|\mathcal{F}_0\right)\right)^{-1}}\\ &= \frac{\mathbb{E}^Q\left(e^{-\int_t^T r_s ds}\mathbb{E}^Q\left(e^{-\int_T^S r_s ds+wY(T,S)}\mid\mathcal{F}_T\right)\mid\mathcal{F}_t\right)}{\mathbb{E}^Q\left(e^{-\int_t^S r_s ds}\,|\mathcal{F}_t\right)}.\end{aligned}$$

The $\mathcal{F}_T$ conditional expectation in this last equation, is also equal to

$$\mathbb{E}^Q\left(e^{-\int_T^S r_s ds+wY(T,S)}\mid\mathcal{F}_T\right) = e^{wY(T,S)}\mathbb{E}^Q\left(e^{-\int_T^S r_s ds}\mid\mathcal{F}_T\right)\,,$$

and, according the Corollary 3.5, we have that

$$\begin{aligned}\mathbb{E}^Q\left(e^{-\int_T^S r_s ds+wY(T,S)}\mid\mathcal{F}_T\right) &= \exp\left(\left(\frac{w}{S-T}-1\right)\left(\int_T^S\varphi(s)ds - A^S(T,S)\right)\right)\\ &\times\exp\left(\left(\frac{w}{S-T}-1\right)\left(X_T(S-T)-\begin{pmatrix}B_1^S(T,S)\\B_2^S(T,S)\end{pmatrix}^\top\begin{pmatrix}\lambda_T^{1Q}\\ \lambda_T^{2Q}\end{pmatrix}\right)\right)\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}^Q\left(e^{-\int_t^S r_s ds}\mid\mathcal{F}_t\right) &= \exp\left(-X_t(S-t)-\int_t^S\varphi(s)ds + A^S(t,S)\right.\\ &\quad\left.+\begin{pmatrix}B_1^S(t,S)\\B_2^S(t,S)\end{pmatrix}^\top\begin{pmatrix}\lambda_t^{1Q}\\ \lambda_t^{2Q}\end{pmatrix}\right).\end{aligned}$$

Using Proposition 2.4 allows us to conclude. □

A.10. Proposition 3.8

Proof. The density of $Y(T,S)$ is retrieved by calculating the Fourier transform of $\varphi^{t,T}(iz)$ as follows:

$$\begin{aligned}f_{Y(T,S)}(y_k) &= \frac{1}{2\pi}\mathcal{F}[\varphi^{t,S}(iz)](y)\\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\varphi^{t,S}(iz)\,e^{-i\,y_k\,z}dz\\ &= \frac{1}{\pi}Re\left(\int_0^{+\infty}\varphi^{t,S}(iz)e^{-i\,y_k\,z}dz\right)\end{aligned}$$

where the last equality comes from the fact that $\varphi^{t,S}(z)$ and $\varphi^{t,T}(-z)$ are complex conjugate. At points $y_k = -\frac{M}{2}\Delta_y + (k-1)\Delta_y$, this last integral is approached with the trapezoid rule

$$\int_a^b h(z)dz = \frac{h(a)+h(b)}{2}\Delta_z + \sum_{k=1}^{M-1} h(a+k\,\Delta_z)\Delta_z$$

and leads to the following estimate for $f_{Y(T,S)}(y_k)$:

$$\begin{aligned}f_{Y(T,S)}(y_k) &\approx \frac{1}{\pi}Re\left(\sum_{j=1}^M \delta_j\varphi^{t,S}(iz_j)e^{-i\,y_k z_j}\Delta_z\right)\\ &\approx \frac{1}{\pi}Re\left(\sum_{j=1}^M \delta_j\varphi^{t,S}(iz_j)(-1)^{j-1}e^{-i\frac{2\pi}{M}(j-1)(k-1)}\Delta_z\right).\end{aligned}$$

References

Acharya, V.V., Pedersen, L.H., 2005. Asset pricing with liquidity risk. J. Financ. Econ. 77 (2), 375–410.

Aït-Sahalia, Y., Cacho-Diaz, J., Laeven, R.J.A., 2015. Modeling financial contagion using mutually exciting jump processes. J. Financ. Econ. 117 (3), 585–606.

Aït-Sahalia, Y., Laeven, R.J.A., Pelizzon, L., 2014. Mutual excitation in eurozone sovereign CDS. J. Econometr. 183, 151–167.

Bacry, E., Delattre, S., Hoffmann, M., Muzy, J.F., 2013. Modelling microstructure noise with mutually exciting point processes. Quant. Finan. 13 (1), 65–77.

Bacry, E., Mastromatteo, I., Muzy, J.F., 2015. Hawkes processes in finance. Market Microstruct. Liq. 1 (1).

Bali, T., 2007. An extreme value approach to estimating interest-rate volatility: pricing implications for interest-rate options. Manag. Sci. 53 (2), 323–339.

Bauwens, L., Hautsch, N., 2009. Handbook of Financial Time Series: Modelling Financial High Frequency Data Using Point Processes. Springer, Berlin.

Bormetti, G., Calcagnile, L.M., Treccani, M., Corsi, F., Marmi, S., Lillo, F., 2015. Modelling systemic price cojumps with Hawkes factor models. Quant. Finan. 15 (7), 1137–1156.

Brigo and Mercurio, 2007. Interest Rate Models — Theory and Practice. Springer Finance

Chavez-Demoulin, V., Davison, A.C., McNeil, A.J., 2005. A point process approach to value-at-risk estimation. Quant. Finan. 5 (2), 227–234.

Chavez-Demoulin, V., McGill, J.A., 2012. High-frequency financial data modeling using Hawkes processes. J. Bank. Financ. 36, 3415–3426.

Chen, L., Lesmond, D.A., Wei, J., 2007. Corporate yield spreads and bond liquidity. J. Finan. 62 (1), 119–149.

Covitz, D., Downing, C., 2007. Liquidity or credit risk? the determinants of very short-term corporate yield spreads. J. Finan. 62 (5), 2303–3232.

Cox, J.C., Ingersoll, J., Ross, S.A., 1985. A theory of the term structure of interest rates. Econometrica 53, 385–407.

Dai, Q., Singleton, K.J., 2000. Specification analysis of affine term structure models. J. Finan. 55, 1943–1978.

Dassios, A., Zhao, H., 2012. Ruin by dynamic contagion claims. Math. Econ. 51 (1), 93–106.

Duffie, D., Kan, R., 1996. A yield-factor model of interest rates. Math. Financ. 6, 379–406.

Eberlein, E., Kluge, W., 2006. Exact pricing formulae for caps and swaptions in a levy term structure model. J. Comput. Financ. 9 (2), 99–125.

Embrechts, P., Liniger, T., Lu, L., 2011. Multivariate Hawkes processes: an application to financial data. J. Appl. Probab. 48 (A), 367–378.

Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modelling Extremal Events for Insurance and Finance. Springer Verlag, Berlin.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987–1007.

Errais, E., Giesecke, K., Goldberg, L.R., 2010. Affine point processes and portfolio credit risk. SIAM J. Financ. Math. 1, 642–665.
Filipović, D., Tappe, S., 2008. Existence of Lévy term structure models. Financ. Stochast. 12 (1), 83–115.
Giot, P., 2005. Market risk models for intraday data. Eur. J. Financ. 11 (4), 309–324.
Hainaut, D., MacGilchrist, R., 2010. An interest rate tree driven by a Lévy process. J. Deriv. 18 (2), 33–45.
Hainaut, D., 2013. A fractal version of the Hull–White interest rate model. Econ. Model. 31, 323–334.
Hainaut, D., 2016. A model for interest rates with clustering effects. Quantitative Finance http://dx.doi.org/10.1080/14697688.2015.1135251. forthcoming.
Hawkes, A., 1971a. Point sprectra of some mutually exciting point processes. J. R. Stat. Soc. B33, 438–443.
Hawkes, A., 1971b. Spectra of some self-exciting and mutually exciting point processes. Biometrika 58, 83–90.
Hawkes, A., Oakes, D., 1974. A cluster representation of a self-exciting process. J. Appl. Probab. 11, 493–503.
Hull, J., White, A., 1990. Pricing interest rate derivatives. Rev. Financ. Stud. 3 (4), 573–592.
Landschoot, A.V., 2008. Determinants of yield spread dynamics: Euro versus us dollar corporate bonds. J. Bank. Financ. 32 (12), 2597–2605.
Longstaff, F.A., 2004. The flight-to-liquidity premium in U.S. treasury bond prices. J. Bus. 77 (3), 511–526.
Mishkin, F.S., 2007. The Economics of Money, Banking, and Financial Markets. Eighth Edition, Pearson International Edition.
Rambaldi, M., Pennesi, P., Lillo, F., 2015. Modeling foreign exchange market activity around macroeconomic news: Hawkes-process approach. Phys. Rev. E. 91 (1), 819–834.
Shen, Y., Siu, T.K., 2013. Pricing bond options under a Markovian regime-switching Hull-White model. Econ. Model. 30, 933–940.
Stabile, G., Torrisi, G.L., 2010. Risk processes with non-stationary Hawkes arrivals. Methodol. Comput. Appl. Probab. 12, 415–429.
Vasicek, O., 1977. An equilibrium characterisation of the term structure. J. Financ. Econ. 5, 177–188.
Wilks, S.S., 1938. The large-sample distribution of the likelihood ratio for testing composite hypotheses. Ann. Math. Stat. 9, 60–62.
Zhang, X., Blanchet, J., Giesecke, K., Glynn, P.W., 2015. Affine point processes: approximation and efficient simulation. Math. Oper. Res. 40 (4), 797–819.
Zhang, X.W., Glynn, P.W., Giesecke, K., 2009. Rare Event Simulation for a Generalized Hawkes Process.
Zhu, L., 2013. Non Linear Hawkes Processes. New York University.
Zhu, L., 2013. Ruin probabilities for risk processes with non-stationary arrivals and subexponential claims. Insurance: Mathematics and Economics 53 (3), 544–550.
Zhu, L., 2014. Limit theorems for a Cox–Ingersoll–Ross process with Hawkes jumps. J. Appl. Probab. 51, 699–712.