"A bivariate Hawkes process for interest rate modeling"

Hainaut, Donatien

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 the level of mutual-excitation.

Document type: Article de périodique (Journal article)

Référence bibliographique
DOI: 10.1016/j.econmod.2016.04.016

Available at:
http://hdl.handle.net/2078/185380
[Downloaded 2019/03/20 at 02:08:16]
A bivariate Hawkes process for interest rate modeling

Donatien Hainaut
ESC Rennes Business School & CREST, France

A R T I C L E   I N F O
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

A B S T R A C T
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 the level 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-
ing a formula for bond pricing is proposed. The Section 4 is about the
interest rate. In order to define this line is as follows:

\[ P_2 = L^2 - \theta_2 B_2, \]

where \( L^2 \) and \( \theta_2 \) are respectively the intercept and the elasticity of the
bond 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:

\[ P^* = L^2 - \theta_2 B^* = L^1 + \theta_1 B^*. \]

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
time, \( t \). According to relation (1), the volume of exchanged bonds at
any given time is hence equal to

\[ B_t = L^2_t - \frac{L^1_t}{\theta_1 + \theta_2}, \]

and the equilibrium bond price at time \( t \), that is inversely propor-
tional to the equilibrium yield is given by

\[ P^*_t = \frac{\theta_1}{\theta_1 + \theta_2} L^1_t - \frac{\theta_2}{\theta_1 + \theta_2} L^2_t \times \frac{1}{r_t}. \]

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:

\[ r_t \propto \frac{\theta_2}{\theta_1 + \theta_2} L^1_t - \frac{\theta_1}{\theta_1 + \theta_2} L^2_t. \]

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.

2. Model

The proposed approach to the analysis of interest rate determi-
nation 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 relation-
ship 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:

\[ P_1 = L^1 + \theta_1 B_1. \]
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.

\[
r_t = \varphi(t) + X_t,
\]

(5)

with

\[
X_t = \alpha_1 L_1^t - \alpha_2 L_2^t
\]

(6)
on a complete probability space \((\Omega, \mathcal{F}, \mathbb{P})\), with a right-continuous information filtration \( \mathcal{F} = (\mathcal{F}_t)_{t \geq 0} \), where \( \mathbb{P} \) 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 \( N_1^t \), \( N_2^t \). From now on, \( L_1^t \) and \( L_2^t \) 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 \( t \):

\[
L_1^t = \sum_{i=1}^{N_1^t} O_1^t,
\]

(7)

\[
L_2^t = \sum_{i=1}^{N_2^t} O_2^t.
\]

(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 \( \alpha_1 \) and \( \alpha_2 \) respectively denotes the permanent impact of sell and buy orders of bonds, the economics theory suggests therefore the following dynamics for \( X_t \):

\[
dX_t = \alpha_1 dL_1^t - \alpha_2 dL_2^t
\]

\[
= \alpha_1 O^1 dN_1^t - \alpha_2 O^2 dN_2^t.
\]

(9)

The order sizes, \( O^1 \) and \( O^2 \), 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 \( \mu_1 = \mathbb{E}(O^1) \), \( \mu_2 = \mathbb{E}(O^2) \), \( \eta_1 = \mathbb{E}((O^1)^2) \), \( \eta_2 = \mathbb{E}((O^2)^2) \). 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.

<table>
<thead>
<tr>
<th>Maturity (q)</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Minimum</th>
<th>Maximum</th>
<th>( \rho(175 \text{ days}) )</th>
<th>( \rho(250 \text{ days}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0217</td>
<td>0.0147</td>
<td>0.0029</td>
<td>0.0545</td>
<td>0.7929</td>
<td>0.6759</td>
</tr>
<tr>
<td>2</td>
<td>0.0229</td>
<td>0.0140</td>
<td>0.0031</td>
<td>0.0548</td>
<td>0.8114</td>
<td>0.7315</td>
</tr>
<tr>
<td>3</td>
<td>0.0245</td>
<td>0.0134</td>
<td>0.0038</td>
<td>0.0540</td>
<td>0.8222</td>
<td>0.7614</td>
</tr>
<tr>
<td>4</td>
<td>0.0261</td>
<td>0.0127</td>
<td>0.0049</td>
<td>0.0528</td>
<td>0.8265</td>
<td>0.7746</td>
</tr>
<tr>
<td>5</td>
<td>0.0277</td>
<td>0.0121</td>
<td>0.0062</td>
<td>0.0519</td>
<td>0.8271</td>
<td>0.7800</td>
</tr>
<tr>
<td>6</td>
<td>0.0292</td>
<td>0.0115</td>
<td>0.0077</td>
<td>0.0513</td>
<td>0.8273</td>
<td>0.7837</td>
</tr>
<tr>
<td>7</td>
<td>0.0305</td>
<td>0.0110</td>
<td>0.0093</td>
<td>0.0510</td>
<td>0.8265</td>
<td>0.7852</td>
</tr>
<tr>
<td>8</td>
<td>0.0317</td>
<td>0.0105</td>
<td>0.0109</td>
<td>0.0508</td>
<td>0.8252</td>
<td>0.7855</td>
</tr>
<tr>
<td>9</td>
<td>0.0327</td>
<td>0.0102</td>
<td>0.0124</td>
<td>0.0507</td>
<td>0.8232</td>
<td>0.7844</td>
</tr>
<tr>
<td>10</td>
<td>0.0337</td>
<td>0.0099</td>
<td>0.0138</td>
<td>0.0509</td>
<td>0.8197</td>
<td>0.7813</td>
</tr>
<tr>
<td>11</td>
<td>0.0353</td>
<td>0.0094</td>
<td>0.0161</td>
<td>0.0513</td>
<td>0.8114</td>
<td>0.7728</td>
</tr>
<tr>
<td>12</td>
<td>0.0370</td>
<td>0.0090</td>
<td>0.0187</td>
<td>0.0516</td>
<td>0.7990</td>
<td>0.7596</td>
</tr>
<tr>
<td>13</td>
<td>0.0380</td>
<td>0.0089</td>
<td>0.0187</td>
<td>0.0526</td>
<td>0.7904</td>
<td>0.7507</td>
</tr>
</tbody>
</table>

---

**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:

$$d\lambda_i^t = \kappa_i \left( \lambda_i -\lambda_i \right) dt + \delta_{1,i} d\lambda_1^t + \delta_{2,i} d\lambda_2^t \quad i = 1, 2,$$

(10)

where $\delta_{ij}$ 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, $\lambda_2^t$, 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 $\gamma_t$ 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 $\Lambda_1^t, \Lambda_2^t$, as discussed later in Section 4.

As shown in Errais et al. (2010), if $f_i = (1_t, \Lambda_i^t)$, the process $(\Lambda_1^t, \Lambda_2^t, j_1^t, j_2^t)$ is a Markov process in the state space $D = (\mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{N})^3$ and its infinitesimal generator for any function $g : D \to \mathbb{R}$ with partial derivatives $g_{\Lambda_1}, g_{\Lambda_2}$ is such that

$$Ag \left( \Lambda_1^t, \Lambda_2^t, j_1^t, j_2^t \right) = \kappa_1 \left( \Lambda_1 - \Lambda_1 \right) g_{\Lambda_1} + \kappa_2 \left( \Lambda_2 - \Lambda_2 \right) g_{\Lambda_2} + \lambda_1^t \int_{-\infty}^t g \left( \Lambda_1 + \delta_{1,1} z_j^t, \Lambda_2 + \delta_{1,2} z_j^t \right) dz_1 \left( z \right) + \lambda_2^t \int_{-\infty}^t g \left( \Lambda_1 + \delta_{2,1} z_j^t, \Lambda_2 + \delta_{1,2} z_j^t \right) dz_2 \left( z \right) - g \left( \Lambda_1^t, \Lambda_2^t, j_1^t, j_2^t \right)$$

(11)

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

$$\mathbb{E} \left( g \left( \Lambda_1^t, \Lambda_2^t, j_1^t, j_2^t \right) \right) = g \left( \Lambda_1^t, \Lambda_2^t, j_1^t, j_2^t \right) + \mathbb{E} \left( \int_t^T Ag \left( \Lambda_1^t, \Lambda_2^t, j_1^t, j_2^t \right) ds \right) \mathbb{F}_t$$

(12)

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

$$\frac{\partial}{\partial T} \mathbb{E} \left( g \left( \Lambda_1^t, \Lambda_2^t, j_1^t, j_2^t \right) \right) = \mathbb{E} \left( Ag \left( \Lambda_1^t, \Lambda_2^t, j_1^t, j_2^t \right) \right) \mathbb{F}_t$$

(13)

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

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

$$m_1(t) = V \left( \frac{1}{\gamma_1} \left( e^{\gamma_1 t} - 1 \right), \frac{1}{\gamma_2} \left( e^{\gamma_2 t} - 1 \right) \right) V^{-1} \left( \kappa_1 c_1, \kappa_2 c_2 \right)$$

$$+ V \left( e^{\gamma_1 t}, e^{\gamma_2 t} \right) V^{-1} \left( \lambda_1^t, \lambda_2^t \right),$$

(14)

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

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

$$V = \begin{pmatrix} \kappa_1 c_1 & \kappa_2 c_2 \\ \frac{1}{\gamma_1} \left( e^{\gamma_1 t} - 1 \right) & \frac{1}{\gamma_2} \left( e^{\gamma_2 t} - 1 \right) \end{pmatrix}$$

(15)

and $\gamma$ is the determinant of $V$ defined by

$$\gamma = \frac{1}{2} \left( \frac{\delta_{1,1} \mu_1 - \kappa_1}{\delta_{1,1} \mu_1 - \kappa_1 - \gamma_1} - \frac{-\delta_{1,1} \mu_2}{\delta_{1,1} \mu_2 - \kappa_2} \right)$$

(16)

and $T$ is the determinant of $V$ defined by

$$T := \delta_{1,2} \mu_2 \sqrt{\left( \frac{\delta_{1,1} \mu_1 - \kappa_1}{\delta_{1,1} \mu_1 - \kappa_1 - \gamma_2} - \frac{-\delta_{1,1} \mu_2}{\delta_{1,1} \mu_2 - \kappa_2} \right)^2 + 4 \delta_{1,2} \delta_{1,2} \mu_1 \mu_2}.$$
The dynamics of \( X_t \) is not mean-reverting and the short term rate, equal to \( \phi(t) + X_t \), may diverge at long term if \( \phi(t) \) is constant. Even if numerical implementations 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 \( \phi(t) \). For example, if we want to keep constant the average short term rate, we can set \( \phi(t) = r_0 - \mathbb{E}(X_t|F_0) \), such as defined by Eq. (6). In this case, the expected rate stays equal to \( r_0 \), whatever the maturity. If we want that expected short term rates revert at long term to \( r_0 \), \( \mathbb{E}(X_t|F_\infty) \), we can set \( \phi(t) = (e^{-at} - 1)r_0 + e^{-at}r_0 - \mathbb{E}(X_t|F_0) \), 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 \( \lambda_i \) by \( V_i(t) = \mathbb{E}((\lambda_i(t))^2) - (\mu_i(t))^2 \) for \( i = 1, 2 \) and their covariance by \( V_3(t) = \mathbb{E}(\lambda_1(t)\lambda_2(t)) \). They are solutions of the following system of ODEs:

\[
\begin{align*}
\frac{d}{dt} \begin{pmatrix} \lambda_1(t) \\ \lambda_2(t) \end{pmatrix} &= \begin{pmatrix}
\beta_{1,1} & \beta_{1,2} \\
\beta_{2,1} & \beta_{2,2}
\end{pmatrix}
\begin{pmatrix} \lambda_1(t) \\ \lambda_2(t) \end{pmatrix} \\
&+ \begin{pmatrix}
2(\beta_{1,2} - \beta_{1,1}) \\
2(\beta_{2,2} - \beta_{2,1})
\end{pmatrix}
\begin{pmatrix} \mu_1(t) \\ \mu_2(t) \end{pmatrix} \\
&\times \begin{pmatrix} V_1(t) \\ V_2(t) \end{pmatrix},
\end{align*}
\]

with initial conditions

\[ V_i(0) = 0 \quad \text{for} \quad i = 1, 2, 3. \]

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.

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 \( \psi_i(\cdot) \) and \( \psi_2(\cdot) \) denote the moment generating functions of \( O^i \) and \( \overline{O}^i \):

\[
\psi_i(w) := \mathbb{E}(e^{w\phi_t}) \quad i = 1, 2.
\]

The moment generating function of \( \phi_t = w_0X_t - w_1 \int_0^T X_s ds + (w_2^	op \begin{pmatrix} \lambda_1^2 \\ \lambda_2^2 \end{pmatrix} ) \) is an affine function of \( X_t \) and of intensities

\[
\mathbb{E}(e^{w_0X_t - w_1 \int_0^T X_s ds + (w_2^	op \begin{pmatrix} \lambda_1^2 \\ \lambda_2^2 \end{pmatrix} ) | F_t})
\]

\[
= \exp\left((w_0 - w_1(T - t))X_t + A(t,T) + \begin{pmatrix} B_1(t,T) \\ B_2(t,T) \end{pmatrix}^\top \begin{pmatrix} \lambda_1^2 \\ \lambda_2^2 \end{pmatrix}\right)
\]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Std. err.</th>
<th>Parameter</th>
<th>Value</th>
<th>Std. err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \kappa_1 )</td>
<td>3.86</td>
<td>0.0187</td>
<td>( \delta_1 )</td>
<td>5619.40</td>
<td>10.3072</td>
</tr>
<tr>
<td>( \kappa_2 )</td>
<td>3.93</td>
<td>0.0196</td>
<td>( \delta_2 )</td>
<td>-4681.94</td>
<td>8.1994</td>
</tr>
<tr>
<td>( \epsilon_1 )</td>
<td>120.00</td>
<td>0.0643</td>
<td>( \delta_{2,1} )</td>
<td>-5486.30</td>
<td>10.1354</td>
</tr>
<tr>
<td>( \epsilon_2 )</td>
<td>134.63</td>
<td>0.0621</td>
<td>( \delta_{2,2} )</td>
<td>4608.52</td>
<td>7.5390</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \rho_1 )</td>
<td>5653.08</td>
<td>6.0047</td>
<td>( \rho_2 )</td>
<td>5576.20</td>
<td>6.0875</td>
</tr>
</tbody>
</table>

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 \( \lambda_1^2 \) and \( \lambda_2^2 \) obtained by log likelihood maximization.
where \( A(t), B_1(t,T) \) and \( B_2(t,T) \) are solutions of a system of ODEs as follows:

\[
\begin{align*}
\frac{d}{dt} B_1(t,T) &= \kappa_1 B_1(t,T) - [\psi_1 (B_1(t,T) \delta_{1,1} + w_0 \alpha_1 - w_1 \alpha_1 (T-t)) + B_2(t,T) \delta_{1,2}) - 1] \\
\frac{d}{dt} B_2(t,T) &= \kappa_2 B_2(t,T) - [\psi_2 (B_1(t,T) \delta_{2,1} - w_0 \alpha_2 + w_1 \alpha_2 (T-t)) + B_2(t,T) \delta_{2,2}) - 1] \\
\frac{d}{dt} A(T) &= -\kappa_1 c_1 B_1(t,T) - \kappa_2 c_2 B_2(t,T)
\end{align*}
\]

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

Notice that for sufficiently large \( \omega_h = 1, \ldots, 4 \), there is no guarantee 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 \((\lambda^1_1, \lambda^1_2, \lambda^2_1, \lambda^2_2)\), 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 local martingale of the form

\[
M_t(\theta_1, \theta_2) := \expleft( (a_1(\theta_1, \theta_2)), a_2(\theta_1, \theta_2) \right) \left( \frac{\lambda^1_1}{\lambda^2_1} \right)
\]

\[
+ (\theta_1, \theta_2) \left( \frac{\lambda^2_1}{\lambda^2_2} \right) - \varphi(\theta_1, \theta_2)t \right)
\]

(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 \((\theta_1, \theta_2) \) fulfill to guarantee that \( M_t(\theta_1, \theta_2) \) is a local martingale.

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

\[
\begin{align*}
\frac{\partial}{\partial t} a_1(\theta_1, \theta_2) \delta_{1,1} - (\psi_1 (a_1(\theta_1, \theta_2) \delta_{1,1} + a_2(\theta_1, \theta_2) \delta_{2,1} + \theta_1)) = 0 \\
\frac{\partial}{\partial t} a_2(\theta_1, \theta_2) \delta_{2,1} - (\psi_2 (a_2(\theta_1, \theta_2) \delta_{2,2} + a_1(\theta_1, \theta_2) \delta_{1,2} + \theta_2)) = 0
\end{align*}
\]

(25)

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

Table 3

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Deviance</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \delta_{1,1} = 0 )</td>
<td>24.08</td>
<td>0.0001%</td>
</tr>
<tr>
<td>( \delta_{1,1} = 0 )</td>
<td>44.16</td>
<td>0.0000%</td>
</tr>
<tr>
<td>( \delta_{2,1} = 0 )</td>
<td>24.05</td>
<td>0.0001%</td>
</tr>
<tr>
<td>( \delta_{2,1} = 0 )</td>
<td>202.88</td>
<td>0.0000%</td>
</tr>
<tr>
<td>( \delta_{1,2} = 0 )</td>
<td>68.21</td>
<td>0.0000%</td>
</tr>
<tr>
<td>( \delta_{1,2} = 0 )</td>
<td>290.24</td>
<td>0.0000%</td>
</tr>
</tbody>
</table>

Table 4

<table>
<thead>
<tr>
<th>Parameters of the Vasicek model, obtained by log likelihood maximization.</th>
<th>Estimates</th>
<th>Std. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>0.014</td>
<td>0.0003</td>
</tr>
<tr>
<td>( \beta )</td>
<td>-0.103</td>
<td>0.0017</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>0.004</td>
<td>0.0001</td>
</tr>
<tr>
<td>Loglik.</td>
<td>19078</td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>-38150</td>
<td></td>
</tr>
</tbody>
</table>

Table 5

| Parameters of a GARCH (1,0) and GARCH (2,0) model for \( \gamma_t \). | Estimate | Std. error | t value | Pr(>|t|) |
|---|---|---|---|---|
| \( \mu \) | 1.652e-06 | 4.543e-06 | 0.364 | 0.716 |
| \( \alpha_0 \) | 4.844e-08 | 1.716e-09 | 28.222 | < 2e-16 *** |
| \( \beta_0 \) | 4.211e-01 | 4.606e-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 $\mu_1 = \frac{1}{\sqrt{\pi}}$, $\mu_2 = -\frac{1}{\sqrt{\pi}}$ and to $\eta_1 = \frac{2}{\sqrt{\pi} \sqrt{\pi}}$. The moment generating functions are given by $\psi_1(z) = \frac{\mu_1}{\sqrt{\pi} \sqrt{\pi}}$ for $z < \rho_1$ and $\psi_2(z) = \frac{\rho_2}{\sqrt{\pi} \sqrt{\pi}}$ 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 $v_i^Q(z)$ under $Q$, are defined by the following parameters:

$$
\rho_1^Q = \rho_1 - (\delta_{1,1} \alpha_1 + \delta_{2,1} \alpha_2 + \theta_1),
\rho_2^Q = \rho_2 + (\alpha_2 \delta_{2,2} + \alpha_1 \delta_{1,2} + \theta_2).
$$

Under $Q^{\delta_1 \delta_2}$, dynamics of intensities are preserved as shown in the next corollary.

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

$$
dN_i^Q = \kappa_i (e^{-\lambda_i^Q} - e^{-\lambda_i^P}) dt + \delta_i^Q dL_i^1 + \delta_i^Q dL_i^2 \quad i = 1, 2.
$$

where the parameters defining the process under $Q$ are

$$
c_i^Q = c_i \psi_i(a_1 \delta_{1,1} + a_2 \delta_{2,1} + \theta_i),
\delta_i^Q = \delta_i \psi_i(a_1 \delta_{1,1} + a_2 \delta_{2,1} + \theta_i),
\gamma_i^Q = \gamma_i \psi_i(a_2 \delta_{2,2} + a_1 \delta_{1,2} + \theta_i)
$$

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 risk neutral measure. The price is denoted by

$$
P(t, T, \lambda^1_i, \lambda^2_i, \lambda^Q_i) = E^Q \left(e^{-\int_t^T X_s ds} \mid F_t\right)
$$

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.**

$$
E^Q \left(e^{-\int_t^T X_s ds} \mid F_t\right) = \exp \left(-X_t(T - t) + A(t, T) + \left(\begin{array}{c}
B_1(t, T) \\
B_2(t, T)
\end{array}\right)^T \lambda^Q_i \lambda^Q_i^T\right)
$$

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

$$
\frac{d}{dt} B_1(t, T) = \kappa_1 B_1(t, T) - \psi_1^Q \left(\begin{array}{c}
B_1(t, T) \delta_{1,1}^Q - \alpha_1(T - t) + B_2(t, T) \delta_{2,1}^Q
\end{array}\right) - 1
$$

$$
\frac{d}{dt} B_2(t, T) = \kappa_2 B_2(t, T) - \psi_2^Q \left(\begin{array}{c}
B_1(t, T) \delta_{1,2}^Q + \alpha_2(T - t) + B_2(t, T) \delta_{2,2}^Q
\end{array}\right) - 1
$$

$$
\frac{d}{dt} A(t, T) = -\kappa_1 C_1 B_1(t, T) - \kappa_2 C_2 B_2(t, T)
$$

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 $L_i^1(dt, dz)$ and $L_i^2(dt, dz)$ and such that

$$
L_i^k = \int_0^\infty \int_{B_i^k} I_{k, k}^i dz dt \quad k = 1, 2.
$$

Furthermore, the expectation of these measures are equal to $E^Q \left(L_i^k(dt, dz) \mid F_t\right) = \lambda_i^Q \eta_i \gamma_i \eta_i^2 dz dt$, for $k = 1, 2$. The next corollary details the infinitesimals dynamics of bond prices:

**Corollary 3.6.** Bond prices, $P(t, T, \lambda_i^1, \lambda_i^2, \lambda_i^Q, \lambda_i^Q)$, are ruled by the following SDE:

$$
dP = P r_t dt - \lambda_i^1 Q \left[\psi_1^Q \left(\begin{array}{c}
B_1(t, T) \delta_{1,1}^Q - \alpha_1(T - t) + B_2(t, T) \delta_{2,1}^Q
\end{array}\right) - 1\right] dt
$$

$$
+ P \int_0^{+\infty} \left[\exp \left(\begin{array}{c}
B_1(t, T) \delta_{1,1}^Q - \alpha_1(T - t) + B_1(t, T) \delta_{2,1}^Q
\end{array}\right) z - 1\right] dt
$$

$$
\times I_{1,1}^Q(dt, dz)
$$

$$
- \lambda_i^2 Q \left[\psi_2^Q \left(\begin{array}{c}
B_1(t, T) \delta_{1,2}^Q + \alpha_2(T - t) + B_2(t, T) \delta_{2,2}^Q
\end{array}\right) - 1\right] dt
$$

$$
+ P \int_0^{+\infty} \left[\exp \left(\begin{array}{c}
B_1(t, T) \delta_{1,2}^Q + \alpha_2(T - t) + B_2(t, T) \delta_{2,2}^Q
\end{array}\right) z - 1\right] dt
$$

$$
\times I_{2,2}^Q(dt, dz)
$$

where $L_1^Q(dt, dz)$ and $L_2^Q(dt, dz)$ 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, $E^Q \left(\frac{dP}{P} \mid F_t\right) = r_t dt$, 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

\[
Y(T, S) := -\frac{1}{T-S} \log P(T, S)
\]

\[
= x_T + \frac{1}{T-S} \left( \int_T^S \varphi(s) ds - A(T, S) \right) - \frac{1}{T-S} \left( B_1(T, S) \right)^T \left( \begin{array}{c} \lambda_1^Q \\ \lambda_2^Q \\ \vdots \\ \lambda_M^Q \end{array} \right)
\]

(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 \( (V(Y(T, S)) = N(S - T)[k - Y(T, S)]) \) 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.

\[
f_j(t, r_t, \lambda_t) = \mathbb{E}^Q \left( e^{-\int_t^T r_s ds} V(Y(T, S)) \mid F_t \right)
\]

(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 \( \exp(-\int_t^T r_s ds) \), because the discount factor is drawn out of Eq. (39), under the forward measure. 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 zero coupon bond \( P(t, S) \), is a martingale and the price of the derivative is equal to

\[
\mathbb{E}^Q \left( e^{-\int_t^T r_s ds} V(Y(T, S)) \mid F_t \right) = P(t, S) \mathbb{E}^Q \left( V(Y(T, S)) \mid F_t \right) = P(t, S) \int_t^\infty V(y) f_{Y(T,S)}(y) dy
\]

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

\[
d^S = \frac{B_t}{B_0} P(0, S) = \left( e^{\int_0^t r_s ds} P(t, S) \right)^{-1} \mathbb{E}^Q \left( e^{\int_t^T r_s ds} \mid F_t \right)
\]

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

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

\[
\varphi^{\Delta x}(w) = \mathbb{E}^Q \left( e^{w Y(T, S)} \mid F_t \right) = \exp \left( \left( \frac{w}{T-S} \right) \int_T^S \varphi(s) ds + w x_T \right) \times \exp \left( A^T(t, T) - A^T(t, S) + \left( \begin{array}{c} \lambda_1^Q \\ \lambda_2^Q \\ \vdots \\ \lambda_M^Q \end{array} \right)^T \left( \lambda_1^Q \\ \lambda_2^Q \\ \vdots \\ \lambda_M^Q \right) \right)
\]

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

\[
\begin{align*}
\frac{d}{dt} B^T_1(t, T) &= \kappa_i B^T_2(t, T) \\
\frac{d}{dt} B^T_2(t, T) &= -\psi_i B^T_1(t, T) - \left( w - (S - t) \right) \alpha_1 + B^T_1(t, T) \delta_1^2 - 1 \\
\frac{d}{dt} B^T_3(t, T) &= -\psi_2 B^T_1(t, T) - \left( w - (S - t) \right) \alpha_2 + B^T_1(t, T) \delta_2^2 - 1 \\
\frac{d}{dt} A^T(t, T) &= -\kappa_1 c_1^2 B^T_1(t, T) - \kappa_2 c_2^2 B^T_2(t, T) \\
\end{align*}
\]

with the terminal conditions \( A^T(T, T) = 1 - \frac{w}{T-S} A^T(T, S), B^T(T, T) = 1 - \frac{w}{T-S} B^T(T, S), B^T(T, T) = 1 - \frac{w}{T-S} B^T(T, S) \).

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 \( \Delta y = \frac{2\pi}{2M} \) be this step of discretization. Let us denote \( \Delta x = \frac{2\pi}{2M} \) and \( z_j = (j - 1)\Delta x \), for \( j = 1 \ldots M \). The values of \( f_{Y(T,S)}(y) \) at points \( y_k = -\frac{2\pi}{M} \Delta y + (k - 1)\Delta y \) are approached by the sum

\[
f_{Y(T,S)}(y_k) \approx \frac{2}{M} \Re \left( \sum_{j=1}^M e^{-2\pi i j y_k} \delta_i \varphi^{\Delta x}(i z_j, r_t, \lambda_t)(-1)^j e^{\frac{-i\pi}{M} (j-1)(k-1)} \right)
\]

(41)

where \( \delta_i = \frac{1}{2} (1_{j=1}) + 1_{j \neq 1} \).

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:

\[
\mathbb{E}^Q \left( e^{-\int_t^T r_s ds} V(Y(T, S)) \mid F_t \right) = P(t, T) \sum_{k=1}^{M+1} V(y_k) f_{Y(T,S)}(y_k) \Delta y
\]

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
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 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 $\frac{1}{\theta_1}$ and $-\frac{1}{\theta_2}$. For this choice of distributions, parameters $\alpha_1$ and $\alpha_2$ 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 $\rho_1$ and $\rho_2$ 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 $\lambda_t^1$ and $\lambda_t^2$ are fitted separately by direct log-likelihood maximization procedures. If daily variations of interest rates are denoted by $\Delta r_i = r_i - r_{i-1}$ for $i = 1$ to $n = 2820$ observations and $\Delta t$ is the length of the time interval, the following two optimization problems are solved numerically to find an estimate of parameters:

$$\begin{align*}
(\theta_1, \theta_2, \alpha_1, \alpha_2, \beta_1, \beta_2) &= \text{arg max } \sum_{i=1}^{n} \log \left( \lambda_t^1 \Delta r_i | \Delta t \geq 0 \right) + \left( 1 - \lambda_t^2 \Delta r_i \right) \mathbb{1}_{\Delta t < 0} \\
(\theta_1, \theta_2, \alpha_1, \alpha_2, \beta_1, \beta_2) &= \text{arg max } \sum_{i=1}^{n} \log \left( \lambda_t^2 \Delta r_i | \Delta t \geq 0 \right) + \left( 1 - \lambda_t^1 \Delta r_i \right) \mathbb{1}_{\Delta t < 0}
\end{align*}$$

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

$$\lambda_t^k = \lambda_{t-1}^k + n_k (\epsilon_k - \lambda_{t-1}) \Delta t + \delta_{k,1} |\Delta r_i| \mathbb{1}_{\Delta t \geq 0} + \delta_{k,2} |\Delta r_i| \mathbb{1}_{\Delta t < 0}, \quad k = 1, 2, i = 1, \ldots, n.$$

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 parameters $\delta_1$ and $\delta_2$ 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 $|\delta_{1,2}| \frac{1}{\theta_1} < \delta_{1,2} \frac{1}{\theta_2}$, 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 $|\delta_{2,1}| \frac{1}{\theta_1} > \delta_{2,1} \frac{1}{\theta_2}$.
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 $\delta_{ij}$'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

$$D = -2 \times (\text{log}(\text{Likelihood full model}) - \text{log}(\text{Likelihood reduced model})).$$

According to the results of Wilks (1938), it is asymptotically
distributed as a $\chi^2$ 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 $k_1t$, $k_2t$ 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:

$$L = \sum_{i=1}^{n} \ln \left[ 1_{\Delta r_i \geq 0} \left( \lambda_1^1 \Delta t + \left( 1 - \left( \lambda_1^1 \Delta t \right) \right) v_1(\Delta r_i) + 1_{\Delta r_i < 0} \left( \lambda_2^2 \Delta t + \left( 1 - \left( \lambda_2^2 \Delta t \right) \right) v_2(\Delta r_i) \right) \right].$$

and is equal to 23 455 whereas the AIC (Akaike Information crite-
rion) 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:

$$dr_t = a(b - r_t)dt + \sigma dW_t \quad (43)$$

where $a$, $b$, $\sigma$ 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. As 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:

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

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.
The variance $s^2_t$ depends both on $D_{t-1}$ and on $s^2_{t-1}$:

$$s^2_t = a_0 + b_1 s^2_{t-1}.$$

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 $k_{1t}$ is most of the time around the frequency of demand orders $k_{2t}$, and they have symmetric patterns of evolution. When $k_{2t}$ is far below $k_{1t}$, 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 $\theta_1$ and $\theta_2$, 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: $\theta_1$ and $\theta_2$ are respectively positive and negative and nearly symmetric till 2013. The second subplot presents the evolution of $x_1(a_1 d_{11} + a_2 d_{21} + \theta_1)$ and $x_2(a_2 d_{22} + a_1 d_{12} + \theta_2)$ that multiply parameters $c_1, c_2$ and $d_{ij}$ under $Q$, as stated in the Corollary 3.4. This graph reveals that parameters driving $k_{1t}$ (resp. $k_{2t}$) are always increased (resp. decreased) under $Q$. And the steepness of the yield curve is directly related to the distance between $x_1(a_1 d_{11} + a_2 d_{21} + \theta_1)$ and $x_2(a_2 d_{22} + a_1 d_{12} + \theta_2)$. Around the credit crunch, yield curves are indeed nearly flat and $x_1(a_1 d_{11} + a_2 d_{21} + \theta_1)$ and $x_2(a_2 d_{22} + a_1 d_{12} + \theta_2)$ 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 $\kappa_1, \kappa_2, C^0_1$ and $C^0_2$ 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 $\kappa_2$, have the same effect on the curve. As the average size of orders are inversely proportional to $\rho_1$ and $\rho_2$, increasing these parameters is equivalent to decrease the average amplitude of variations of the interest rate. Then higher $\rho_1$ or $\rho_2$ respectively lowers or raises the steepness of the curve.

Fig. 7 presents simulated sample paths for $r_t$, $\lambda_t^1$ and $\lambda_t^2$ 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, $V(2,3)$ 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 $\gamma_{\text{max}} = 0.10$. 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 $\rho_1$ or $\rho_2$ 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 $g^i = \lambda^i$ for $i = 1, 2$. According to Eqs. (11) and (12), their expectations are such that

$$E(Aq^1) = \kappa_1 \left( c_1 - E(\lambda^1) \right) + E(\lambda^1) \int_{-\infty}^{\infty} \delta_{1,2} d\nu_1(z) + E(\lambda^2) \int_{-\infty}^{\infty} \delta_{1,2} d\nu_2(z)$$

$$\times \left( \delta_{2,1} \delta_{1,2} - \delta_{2,2} \delta_{1,1} \right)$$

$$= \kappa_1 \left( c_1 - E(\lambda^1) \right) + E(\lambda^1) \delta_{1,1} \mu_1 + E(\lambda^2) \delta_{1,2} \mu_2$$

$$E(Aq^2) = \kappa_2 \left( c_2 - E(\lambda^2) \right) + E(\lambda^1) \int_{-\infty}^{\infty} \delta_{2,1} d\nu_1(z) + E(\lambda^2) \int_{-\infty}^{\infty} \delta_{2,2} d\nu_2(z)$$

$$\times \left( \delta_{2,1} \delta_{1,2} - \delta_{2,2} \delta_{1,1} \right)$$

$$= \kappa_2 \left( c_2 - E(\lambda^2) \right) + E(\lambda^1) \delta_{2,1} \mu_1 + E(\lambda^2) \delta_{2,2} \mu_2$$

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.

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

(44)

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

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

Eigenvectors cancel the determinant of the following matrix:

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

and are solutions of the following second order equation:

$$\gamma^2 - \gamma \left( \delta_{2,1} \mu_1 - \kappa_1 \right) - \left( \delta_{1,1} \mu_1 - \kappa_1 \right) \delta_{2,2} \mu_2 - \kappa_2 = 0.$$
Roots of this last equation are $\gamma_1$ and $\gamma_2$, as defined by Eq. (15). One way to find an eigenvector is to note that it must be orthogonal to each row of the following matrix:

$$
\left( \begin{array}{cc}
(\delta_{1,1} \mu_1 - \kappa_1) - \gamma & \delta_{1,2} \mu_2 \\
\delta_{2,1} \mu_1 & (\delta_{2,2} \mu_2 - \kappa_2) - \gamma
\end{array} \right) \left( \begin{array}{c}
v_1 \\
v_2
\end{array} \right) = 0,
$$

then necessarily,

$$
\left( \begin{array}{c}
v'_1 \\
v'_2
\end{array} \right) = \left( \begin{array}{c}
-\delta_{1,2} \mu_2 \\
(\delta_{1,1} \mu_1 - \kappa_1) - \gamma
\end{array} \right) \quad \text{for } i = 1,2.
$$

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

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

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

$$
\left( \begin{array}{c}
u_1 \\
u_2
\end{array} \right) = V^{-1} \left( \begin{array}{c}
m_1 \\
m_2
\end{array} \right).
$$

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

$$
\frac{\partial}{\partial t} \left( \begin{array}{c}
u_1 \\
u_2
\end{array} \right) = V^{-1} \left( \begin{array}{c}
\kappa_1 c_1 \\
\kappa_2 c_2
\end{array} \right) + \left( \begin{array}{cc}
\gamma_1 & 0 \\
0 & \gamma_2
\end{array} \right) \left( \begin{array}{c}
u_1 \\
u_2
\end{array} \right) \quad \text{(45)}
$$

And introducing the following notations:

$$
V^{-1} \left( \begin{array}{c}
\kappa_1 c_1 \\
\kappa_2 c_2
\end{array} \right) = \left( \begin{array}{c}
f_1 \\
f_2
\end{array} \right),
$$

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

$$
u_1(t) = \frac{f_1}{\gamma_1} (e^{\gamma_1 t} - 1) + d_1 e^{\gamma_1 t} \\
u_2(t) = \frac{f_2}{\gamma_2} (e^{\gamma_2 t} - 1) + d_2 e^{\gamma_2 t}
$$

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

$$
\left( \begin{array}{c}
u_1 \\
u_2
\end{array} \right) = \left( \begin{array}{cc}
\frac{1}{\gamma_1} (e^{\gamma_1 t} - 1) & 0 \\
0 & \frac{1}{\gamma_2} (e^{\gamma_2 t} - 1)
\end{array} \right) V^{-1} \left( \begin{array}{c}
\kappa_1 c_1 \\
\kappa_2 c_2
\end{array} \right) + \left( \begin{array}{cc}
e^{\gamma_1 t} & 0 \\
0 & e^{\gamma_2 t}
\end{array} \right) V^{-1} \left( \begin{array}{c}
\lambda_1 \\
\lambda_2
\end{array} \right).
$$

Expressions (Eq. (14)) for $m_1, m_2$ are inferred from this last relation.
A.4. Proposition 2.4

Proof. Let us define 
\[ f(t, \lambda^1_1, \lambda^1_2, \lambda^2_1, \lambda^2_2) := \left( \begin{array}{c} w_{01} \lambda^1_1 - \lambda^1_2 \lambda^2_1 - \lambda^1_2 \lambda^2_2 \\ \int_{-\infty}^{0} e^{\lambda^1_1 s} dS(t) \\ w_{02} \lambda^2_1 - \lambda^1_2 \lambda^2_1 - \lambda^1_2 \lambda^2_2 \\ \int_{-\infty}^{0} e^{\lambda^2_1 s} dS(t) \end{array} \right), \]
and let \( f_1, f_2, \) and \( f_3 \) 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 integral integro-differential equation:

\[ \begin{align*}
&w_1 (\alpha_1 \lambda^1_1 - \alpha_2 \lambda^2_1) = f_1 + \kappa_1 (c_1 - \lambda^1_1) f_1 + \kappa_2 (c_2 - \lambda^2_2) f_2 + \lambda^1_1 \int_{t=0}^{\infty} f(t, \lambda^1_1 + \alpha_1, \lambda^1_2 + \alpha_2, \lambda^2_1, \lambda^2_2) dt + f_1 - f(t, \lambda^1_1 + \alpha_1, \lambda^1_2, \lambda^2_1, \lambda^2_2) d\lambda^1_1 \\
&+ \lambda^2_1 \int_{t=0}^{\infty} f(t, \lambda^1_1, \lambda^1_2 + \alpha_1, \lambda^2_1 + \alpha_2, \lambda^2_2) dt + f_2 - f(t, \lambda^1_1, \lambda^1_2, \lambda^2_1 + \alpha_1, \lambda^2_2, \lambda^2_2) d\lambda^2_1.
\end{align*} \]

(47)

(48)

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

\[ dY_t = a_0 \kappa_1 \left( c_1 - \lambda^1_1 \right) dt + a_2 \kappa_2 \left( c_2 - \lambda^2_2 \right) dt + (a_1 \alpha_1 + a_2 \alpha_2) d\lambda^1_1 + (a_2 \alpha_2 + a_1 \alpha_1 + \theta_1) d\lambda^2_2 - \varphi(\theta_1, \theta_2) d\lambda^1_1. \]

In the remainder of this proof, the random measure \( \Omega' \) is noted \( \chi' \) and is such that \( \Omega' = \int_{i=1}^{\infty} \chi'(dz) \) for \( i = 1, 2 \).

Applying Itô’s lemma for semi-martingales to \( M_1 \) leads to the next relation:

\[ \begin{align*}
dM_1 &= M_1 dY_t + \frac{1}{2} M_1 d[Y_1, Y_1]_t \\
&+ M_1 \int_{t=0}^{\infty} \left( e^{(s \alpha_1 + s \alpha_2 + s \theta_1)} - 1 \right) \chi'(dz) dN_t^1 \\
&+ M_2 \int_{t=0}^{\infty} \left( e^{(s \alpha_2 + s \alpha_2 + s \theta_2)} - 1 \right) \chi'(dz) dN_t^2
\end{align*} \]

or equal to

\[ \begin{align*}
dM_1 &= M_1 (a_0 \kappa_1 \left( c_1 - \lambda^1_1 \right) + a_2 \kappa_2 \left( c_2 - \lambda^2_2 \right) - \varphi(\theta_1, \theta_2)) dt \\
&- M_1 \kappa_1 \left( c_1 - \lambda^1_1 \right) \chi'(dz) dN_t^1 \\
&- M_2 \kappa_2 \left( c_2 - \lambda^2_2 \right) \chi'(dz) dN_t^2 \\
&+ M_1 \int_{t=0}^{\infty} \left( e^{(s \alpha_1 + s \alpha_2 + s \theta_1)} - 1 \right) \chi'(dz) dN_t^1 \\
&+ M_2 \int_{t=0}^{\infty} \left( e^{(s \alpha_2 + s \alpha_2 + s \theta_2)} - 1 \right) \chi'(dz) dN_t^2.
\end{align*} \]

Since the integrals with respect to \( \chi'(dz) dN_t^i = \lambda^i_1 \lambda^i_2 d\lambda^i_1 d\lambda^i_2 \) are local martingales, \( M_1 \) is also a local martingale if and only if the following relations hold:

\[ \begin{align*}
a_1(\theta_1, \theta_2) \kappa_1 c_1 + a_2(\theta_1, \theta_2) \kappa_2 c_2 - \varphi(\theta_1, \theta_2) &= 0 \\
a_1(\theta_1, \theta_2) \kappa_1 c_1 + a_2(\theta_1, \theta_2) \kappa_2 c_2 - \varphi(\theta_1, \theta_2) &= 0
\end{align*} \]

and these conditions are equivalent Eqs. (25) and (26). \( \square \)

A.6. Proposition 3.2

Proof. If \( Y_t \) is the exponent of \( M_0 \), as defined by Eq. (51), the moment generating function of \( X_T \) under the risk neutral is then equal to

\[ \mathbb{E}^\mathbb{Q} \left( e^{wX_T} | \mathcal{F}_T \right) = \mathbb{E} \left( e^{wY_T+wX_T} | \mathcal{F}_T \right) = e^{-wY_T} \mathbb{E} \left( e^{wX_T+wX_T} | \mathcal{F}_T \right). \]

If \( f(t, \lambda^1_1, \lambda^1_2, \lambda^2_1, \lambda^2_2) \) denotes \( e^{(w^1 + w^2) \lambda_T} \), according to the Itô’s lemma, it solves the next equation

\[ \begin{align*}
0 &= \kappa_1 (c_1 - \lambda^1_1) f_1 + \kappa_2 (c_2 - \lambda^2_2) f_2 + \lambda^1_1 \int_{t=0}^{\infty} f(t, \lambda^1_1 + \alpha_1, \lambda^1_2 + \alpha_2, \lambda^2_1, \lambda^2_2) dt + f_1 - f(t, \lambda^1_1 + \alpha_1, \lambda^1_2, \lambda^2_1, \lambda^2_2) d\lambda^1_1 \\
&+ \lambda^2_1 \int_{t=0}^{\infty} f(t, \lambda^1_1, \lambda^1_2 + \alpha_1, \lambda^2_1 + \alpha_2, \lambda^2_2) dt + f_2 - f(t, \lambda^1_1, \lambda^1_2, \lambda^2_1 + \alpha_1, \lambda^2_2, \lambda^2_2) d\lambda^2_1.
\end{align*} \]

(50)

(52)
where \(f, f_1, f_2\) are the partial derivatives of \(f(.)\) with respect to time and intensities. Furthermore given that

\[
Y_T + wX_T = a_1 \left( \lambda_1^1 - \kappa_1 c_1 T \right) + a_2 \left( \lambda_2^2 - \kappa_2 c_2 T \right)
+ \left( \theta_1 + \alpha_1 w \right) L_1^1 + \left( \theta_2 + \alpha_2 w \right) L_2^2.
\]

(53)

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

\[
f \left( T, \lambda_1^1, L_1^1, \lambda_2^2, L_2^2 \right) = \exp \left( \theta_1 + \alpha_1 w L_1^1 + \left( \theta_2 + \alpha_2 w \right) L_2^2 \right)
+ a_1 \left( \lambda_1^1 - \kappa_1 c_1 T \right) + a_2 \left( \lambda_2^2 - \kappa_2 c_2 T \right).\]

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

\[
f = \exp \left( A(tT)^\top \left( \psi_1(a_1 \delta_1 + a_2 \delta_2 + \theta_1) \right) + C(tT)^\top \left( L_1 \right) \right)
\]

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

\[
f_{t_1} = \psi_1 B_1 f, \quad f_{w} = \psi_2 B_2 f
\]

where \(\psi_1\) and \(\psi_2\) abusively denote \(\psi_1(a_1 \delta_1 + a_2 \delta_2 + \theta_1)\) and \(\psi_2(a_2 \delta_2 + \alpha_1 \delta_1 + \theta_2)\). Integrands in Eq. (52) are equal to

\[
f \left( (\lambda_1^1 + \delta_1 z, L_1^1 + z(1), \lambda_2^2 + \delta_2 z, L_2^2) - f \left( \lambda_1^1, L_1^1, \lambda_2^2, L_2^2 \right) \right)
= \exp \left( (B_1 \psi_1 \delta_1 + \psi_2 \delta_2 + C_1 + B_2 \psi_2 \delta_2) z - 1 \right),
\]

(55)

where \(A, B_1, B_2\) and \(C_1, C_2\):

\[
0 = \frac{\partial}{\partial t} B_1 - \kappa_1 B_1 + \left[ \frac{1}{\psi_1} \psi_1 \left( B_1 \psi_1 \delta_1 + \psi_2 \delta_2 + C_1 + B_2 \psi_2 \delta_2 \right) - 1 \right],
\]

\[
0 = \frac{\partial}{\partial t} B_2 - \kappa_2 B_2 + \left[ \frac{1}{\psi_2} \psi_2 \left( B_1 \psi_1 \delta_1 + \psi_2 \delta_2 + C_1 + B_2 \psi_2 \delta_2 \right) - 1 \right],
\]

\[
0 = \frac{\partial}{\partial t} A + \psi_1 \delta_1 C_1 + \psi_2 \delta_2 C_2 B_2,
\]

\[
0 = \frac{\partial}{\partial t} C_1, \quad 0 = \frac{\partial}{\partial t} C_2
\]

with the terminal conditions:

\[
A(T, T) = -a_1 \delta_1 c_1 T - a_2 \delta_2 c_2 T
\]

\[
B_1(T, T) = \frac{\rho_1}{\rho_1 - \rho_1 - \rho_1}, \quad B_2(T, T) = \frac{\rho_2}{\rho_2 - \rho_2 - \rho_2},
\]

\[
C_1(T, T) = \left( \theta_1 + \alpha_1 w \right), \quad C_2(T, T) = \left( \theta_2 + \alpha_2 w \right)
\]

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

\[
E^Q \left( e^{wX_t|\mathcal{F}_t} \right) = e^{-\gamma E^Q \left( e^{r + wX_t|\mathcal{F}_t} \right)}
= \exp \left( A + a_1 \kappa_1 c_1 t + a_2 \kappa_2 c_2 t + (B_1 \psi_1 - a_1) \lambda_1^1 \right)
+ (B_2 \psi_2 - a_2) \lambda_2^2 + wX_t\)
\]

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:

\[
A' := A + a_1 \kappa_1 c_1 t + a_2 \kappa_2 c_2 t,
B_1' := B_1 - \frac{\alpha_1}{\psi_1},
B_2' := B_2 - \frac{\alpha_2}{\psi_2}
\]

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:

\[
\begin{align*}
\kappa_1 \frac{\psi_1}{\psi_1} &= \left( 1 - \frac{1}{\kappa_1} \right), \\
\kappa_2 \frac{\psi_2}{\psi_2} &= \left( 1 - \frac{1}{\kappa_2} \right)
\end{align*}
\]

(55)

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

\[
0 = \frac{\partial}{\partial t} B_1' - \kappa_1 B_1' + \left[ \frac{1}{\psi_1} \psi_1 \left( B_1' \psi_1 \delta_1 + (\theta_1 + \delta_1 a_1 + \delta_2 a_2) \right) - 1 \right],
\]

\[
0 = \frac{\partial}{\partial t} B_2' - \kappa_2 B_2' + \left[ \frac{1}{\psi_2} \psi_2 \left( B_1' \psi_1 \delta_2 + (\theta_2 + \delta_1 a_1 + \delta_2 a_2) \right) - 1 \right],
\]

\[
0 = \frac{\partial}{\partial t} A' + \psi_1 \kappa_1 c_1 B_1' + \psi_2 \kappa_2 c_2 B_2'.
\]

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

\[
E^Q \left( e^{wX_t|\mathcal{F}_t} \right) = \exp \left( wX_t + A'(t, T) + \left( B_1' B_2' \right)^\top \left( \lambda_1^{1,0}, \lambda_2^{2,0} \right) \right)
\]

(56)

where \(A', B'_1, B'_2\) solve a system, identical to the one of Proposition 2.4.

\(\square\)

A.7. Corollary 3.3

**Proof.** If we denote \(\beta_1 = \delta_1 a_1 + \delta_2 a_2 + \theta_1\), by construction the moment generating function of sell orders, under the risk neutral measure is provided by the following ratio:

\[
\psi_1^Q (z) = \frac{\psi_1 (z + \beta_1)}{\psi_1 (\beta_1)} = \frac{\rho_1}{\rho_1 - z - \beta_1}
\]

(57)
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 Itô’s lemma for semi-martingales, $P \left( t, \lambda_1^{(Q)} J_1^{(Q)}, \lambda_2^{(Q)} J_2^{(Q)} \right)$ is such that

$$
dP = P_t + \int_{t}^{+\infty} P \left( \lambda_1^{(Q)} J_1^{(Q)} + \left( z, 1 \right)^T \lambda_2^{(Q)} J_2^{(Q)} \right) P_{\lambda \Delta t} dt
+ \int_{t}^{+\infty} P \left( \lambda_1^{(Q)} J_1^{(Q)} \lambda_2^{(Q)} J_2^{(Q)} \right) 1^{Q \lambda \Delta t} dt, dz
+ \int_{t}^{+\infty} P \left( \lambda_1^{(Q)} J_1^{(Q)} \lambda_2^{(Q)} J_2^{(Q)} \right) 1^{Q \lambda \Delta t} dt, dz,
$$

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:

$$
\mathcal{E}^Q \left( e^{W(T,S)} | \mathcal{F}_T \right) = \mathcal{E}^Q \left( e^{\int_0^T W(t) dt} | \mathcal{F}_0 \right)^{-1} \mathcal{E}^Q \left( e^{W(T,S)} | \mathcal{F}_T \right)
= \mathcal{E}^Q \left( e^{-\frac{1}{2} \sigma^2 W(T,S)} | \mathcal{F}_T \right).
$$

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

$$
\mathcal{E}^Q \left( e^{-\frac{1}{2} \sigma^2 W(T,S)} | \mathcal{F}_T \right) = e^{W(T,S)} \mathcal{E}^Q \left( e^{-\frac{1}{2} \sigma^2 W(T,S)} | \mathcal{F}_T \right),
$$

and, according the Corollary 3.5, we have that

$$
\mathcal{E}^Q \left( e^{-\frac{1}{2} \sigma^2 W(T,S)} | \mathcal{F}_T \right) = \exp \left( \frac{W}{S - T} - 1 \right) \left( \int_0^T \varphi(s) ds - A^Q(T,S) \right)
\times \exp \left( \frac{W}{S - T} - 1 \right) X_T (S - T) - \left( B_1^Q(T,S) \right)^T \lambda_1^{(Q)} + \left( B_2^Q(T,S) \right)^T \lambda_2^{(Q)}
$$

and

$$
\mathcal{E}^Q \left( e^{-\frac{1}{2} \sigma^2 W(t)} | \mathcal{F}_T \right) = \exp \left( -\frac{1}{2} \sigma^2 (S - t) - \int_t^S \varphi(s) ds + A^Q(t,S) \right)
+ \left( B_1^Q(T,S) \right)^T \lambda_1^{(Q)} + \left( B_2^Q(T,S) \right)^T \lambda_2^{(Q)}.
$$

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 $\varphi^T(z)$ as follows:

$$
f_{Y(T,S)}(y) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \varphi^T(z) \varphi(z) dz
= \frac{1}{\pi} \Re \left( \int_0^{+\infty} \varphi^T(z) e^{-i y z} dz \right)
$$

where the last equality comes from the fact that $\varphi^T(z)$ and $\varphi^T(-z)$ 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

$$
\int_0^b h(z) dz = \frac{h(a) + h(b)}{2} \Delta z + \sum_{k=1}^{M-1} h(a + k \Delta z) \Delta z
$$

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

$$
f_{Y(T,S)}(y_k) \approx \frac{1}{\pi} \Re \left( \sum_{j=1}^{M} \varphi^T(z_j) e^{-i y_j \Delta z} \right)
\approx \frac{1}{\pi} \Re \left( \sum_{j=1}^{M} \varphi^T(z_j) \left( -1 \right)^{j-1} e^{-i \frac{\pi}{2} (j-1)(k-1)} \Delta z \right).
$$

References


