A multifactor self-exciting jump diffusion approach for modelling the clustering of jumps in equity returns

Donatien Hainaut*
Univ. Catholique de Louvain

Franck Moraux†
Univ. Rennes 1 and CREM

June, 1 2016

Abstract

This paper introduces a new jump diffusion process where the occurrence and the size of past jumps have an impact on both the instantaneous and the long term propensities of observing a jump instantaneously. Here, the intensity of jump arrival is a *multifactor* self-excited process whereas the jump size is a double exponential random variable. This specification capture many dynamic features of asset returns; it can for instance handle with the jump clustering effects explored by Ait-Sahalia et al. (2015). Moreover, it remains analytically tractable, as we can prove that these multifactor self-excited processes are similar to single factor processes whose kernel function is the sum of two exponential functions. We can derive various closed and semi-closed form expressions for the mean and the variance of the intensity as well as for the moment generating of log returns. We also find a class of changes of measure that preserves the dynamics of the process under the risk neutral measure. To motivate empirically the multifactor model, we calibrate the model by a peak over threshold approach and filter state variables by sequential Monte Carlo algorithm. We also investigate if self-excitation is induced by positive, negative or both jumps. So as to illustrate the applicability of our modeling for derivatives, we next evaluate European options and analyze the sensitivity of implied volatilities to parameters and factors.

*Postal address: ISBA, UCL,Voie du Roman Pays 20, B-1348 Louvain-la-Neuve (Belgique). E-mail to: donatien.hainaut@gmail.com.

†Postal address: IGR-IAE Rennes, 11 rue Jean Macé, 35000 Rennes (France). E-mail to: franck.moraux@univ-rennes1.fr.

A MULTIFACTOR SELF-EXCITING JUMP DIFFUSION APPROACH FOR MODELLING THE CLUSTERING OF JUMPS IN EQUITY RETURNS

Abstract: This paper introduces a new jump diffusion process where the occurrence and the size of past jumps have an impact on both the instantaneous and the long term propensities of observing a jump instantaneously. Here, the intensity of jump arrival is a *multifactor* self-excited process whereas the jump size is a double exponential random variable. This specification capture many dynamic features of asset returns; it can for instance handle with the jump clustering effects explored by Ait-Sahalia et al. (2015). Moreover, it remains analytically tractable, as we can prove that these multifactor self-excited processes are similar to single factor processes whose kernel function is the sum of two exponential functions. We can derive various closed and semi-closed form expressions for the mean and the variance of the intensity as well as for the moment generating of log returns. We also find a class of changes of measure that preserves the dynamics of the process under the risk neutral measure. To motivate empirically the multifactor model, we calibrate the model by a peak over threshold approach and filter state variables by sequential Monte Carlo algorithm. We also investigate if self-excitation is induced by positive, negative or both jumps. So as to illustrate the applicability of our modeling for derivatives, we next evaluate European options and analyze the sensitivity of implied volatilities to parameters and factors.

JEL Codes: G10, G11.

1 Introduction

The propensity of price jumps to cluster over several days has been recently documented in the financial markets[1]. The phenomenon has been specifically studied and evidenced by Yu (2004) and by Maheu and McCurty (2004) who consider respectively DJIA returns and a list of individual stock returns. More recently, Ait-Sahalia et al (2015) explore international equity market indices on a daily basis and conclude that *"Jump clustering in time is a strong effect in the data"*. They also notice that from *"mid-September to mid-November 2008, the US stock market jumped by more than 5% on 16 separate days"* (see footnote 2, page 587).

This jump clustering phenomenon dramatically questions the traditional understanding of equity market dynamics, as well as the common practices, the routines and the needs for relevant tools of managers, academics and regulators[2]... With no surprise, a recent strand of the literature makes significant efforts to understand the cause of the clustering of jumps, to develop quantitative methods to model this phenomenon[3] and to investigate implications of it for asset pricing, option pricing and risk management. The why we observe some jumps in equity prices has been especially questioned and investigated with high-frequency data and there is a broad consensus in this literature that news releases impact a lot at both the individual and market levels (see e.g. Lee and Mykland (2008) and Evans (2011) for additional evidences)[4]. More generally, it is widely believed that the information flow matters (see Rangel 2011, Lee 2012 and Fulop et al. 2015 who even advocate the use of a Bayesian learning approach). As noticed by Maheu and McCurdy (2004), this flow may explain some clustering effects (*"Like the information process itself, jumps tend to be clustered together"*). Beyond that, *"market crashes can be realized in a series of jumps over a short period"*.

A recent, natural[5] and endogenous way to capture clustering of jumps in the daily

[1]The clustering of jumps is a very common feature in high-frequency data. But we mainly keep this very rich issue and literature out of the scope of the present research. Interested readers may nevertheless find some references below.

[2]Portfolio managers may be, e.g., highly interested in estimates of how often unexpected jumps can occur.

[3]A rich strand of the literature, out the scope of this article, aims at developping statistical tests capable to detect the jumps. The core challenge of this literature is to disentangle extreme realizations of a continuous process and jumps.

[4]Lee and Mykland (2008) conclude that the individual stock jumps are linked to company-specific news events (such that scheduled earnings announcements but also unscheduled news), while the "*S&P 500 Index jumps are associated with general market news announcement*". More recently, Lee (2012) goes a step further by analyzing the predictability of jumps in individual stock returns, using both macroeconomic and firm-specific news releases.

[5]It is worth noticing that many stochastic processes commonly used for modeling jumps are unable to capture the jump clustering effect. Lévy processes, for instance, are very useful to accommodate the skewness and the excess kurtosis of financial security returns, but they are unable to deal with the jump clustering. They are indeed Markovian and have independent increments (*see Schoutens (2003)*). In addition, the clustering of jumps has been 'latent' in many research on daily dynamics far before the

price dynamics is to use self-exciting point processes where the jump arrival intensity at a given point in time (which is the probability to observe a jump instantaneously) depends on the number and sometimes the size of jumps the price of the asset experiments before. This approach is linked to the Hawkes self-exciting processes (see Hawkes (1971a, b) and Hawkes and Oakes (1974))[6]. In the most common and simplest specification, the jump arrival intensity process is persistent and it suddenly increases as soon as a jump occurs in the asset price. Moreover, the jump influence on the intensity does not depend on its size and it decays over time more or less rapidly according to a kernel function. Different types of Hawkes self-exciting processes have been recently introduced for modelling the daily dynamics of financial assets. Aït-Sahalia et al. (2014, 2015) develop a multidimensional setting with self and mutual excitations in order to question whether some jumps result from some contagion effects. The jump arrival intensity processes depend there on the number of realized jumps, not associated sizes. By so doing, Ait-Sahalia et al (2014, 2015) can disentangle the time excitation and the space excitation and they provide evidence that the mutual excitation matters to explain the clustering of jump. Carr and Wu (2016), Chen and Poon (2013) and Fulop et al. (2015) rather develop mono-asset settings where the role of the jump size and, in particular, the negative jumps associated to bad news is key to evaluate asset risk, option prices and variance risk premia.

In this paper, we partly follow previous contributions by exploring whether the number and the size of past jumps can increase the probability to observe a jump in the next future. But we go a step further by introducing a *multifactor* self-excited jump diffusion (SEJD) process. Here the self-exciting contemporaneous jump arrival intensity reverts towards a long-run mean intensity which itself follows a self-exciting Hawkes process. Our modelling approach is multifactor because the jumps are led by two dependent states variables: the contemporaneous jump intensity and its long-run mean level[7]. The other distinctive features of our model are that the size of the realized jumps can impact the two variables and that all past jumps can impact. This is different from Aït-Sahalia et

advent of Hawkes self-exciting process, because jump clustering is a characteristic of any jump model equipped with a time-varying and persistent conditional intensity (see Chan and Maheu (2002), for a typical example). For the sake of modelling daily returns, a discrete jump component has early been included in the price dynamics to accomodate the infrequent large price movements (see Press (1967), Merton (1976)), but the need for a time-varying jump intensity has been recognized for years. Bates (2000) develops a continuous-time setting where the jump intensity depends on the level of a stochastic volatility (see also Andersenet al. (2002), Pan (2002) and Eraker (2004)). Duffie et al. (2000) generalize the approach by assuming that the jump intensity may be an affine function of a latent variable. An far simpler way to allow the jump arrival intensity to vary over time is to add some dummy variables for days, as Das (2002) did for modelling jumps in interest rates, or for the week-end (see Fortune (1999)).

[6]This set of point processes is also commonly used for modelling 'high-frequency' data. Engle and Russell (1998) point out the fundamental role of Hawkes self-exciting processes in the modelling of duration between two transactions. Readers interested by high-frequency applications may consult Giot (2005), Hewlett (2006), Bowsher (2007), Large (2007), Chavez-Demoulin and McGill (2012), Bacry et al (2013) and Da Fonseca and Zantour (2014) for more recent contributions. Bauwens et Hautsch (2009) offer an overview.

[7]Hence, this bifactor approach should not be confused with the one of Aït-Sahalia et al. (2015 a,b) where the bifactor feature refers to the number of financial assets simultaneously taken in account.

al. (2014, 2015) where the size of the realized jumps does not matter and from Carr and Wu (2016) and others who take in account only negative ones. The main intuition of our specification is that the occurrence and the size of past realized jumps can influence not only the instantaneous probability of observing a new jump (as usual) but also its long-run mean level, that is the level to which the jump intensity reverts. In our model, both state variables are impacted by the previous occurrences of shocks and their realized sizes in absolute terms. Several jump distributions are tested: positive exponential, negative exponential, double exponential and binomial laws. This means that our model nests the dynamics of Kou (2002) as a special case. Note also that we work with this distribution for its flexibility, but most of our conclusions apply to other size distributions.

The contributions of this paper are multiple and various. First and mainly, we introduce a new multifactor self-exciting Hawkes process that can investigate a possible impact of price jumps on the long-run mean toward which the jump arrival intensity reverts. We provide a number key properties of this new process and demonstrate, among other things, that the jump arrival intensity associated to our multifactor specification may be viewed as the intensity of an single factor self-exciting equipped with a complex kernel function. In theory, our model captures then a larger spectrum of dynamics than common self-exciting Hawkes processes with exponentially decreasing kernels Moreover, our model is highly tractable from a 'stochastic calculus' perspective[8]. We can indeed obtain analytical expressions for the first two centered moments of state variables and a semi closed form expressions for their moment generating functions. We show that the expectation and variance of the jump arrival intensity also depends upon a sum of exponential functions. For the sake of pricing financial and derivatives contracts, we identify a family of changes of measure that preserves the structure of our specification and we use this to define a risk neutral measure. We provide analytical pricing formulas for realized variances. For calibration purposes, we develop an assymmetric peak over threshold procedure. Next, a sequential Monte Carlo procedure is used to filter accurately the hidden states variables and to appraise the likelihood. We explore the empirical performance of our model on daily returns of various financial time series collected over a period of ten years (S&P 500 index). Our analysis also determines which is the most appropriate distributions for jumps, among the positive, negative, double exponential and binomial laws. And the results reveal that the self-excitation is also induced by positive shocks, and not only by negative ones as suggested by Carr & Wu (2016). Next options are priced with the recourse to a Fast Fourier Transform technique to compute the probability density of the asset return from the moment generating function of log return.

The paper proceeds as follows. Section 2 presents the continuous time framework. Section 3 discusses calibration issues and contains an empirical analysis. Section 4 illustrates

[8]Consequently, our specification represents an interesting compromise between the pure Hawkes process whose exponentially decreasing is not always supported by financial data (see Embrechts et al. (2011) or Bacry et al. (2013)) and more intricate self-exciting processes, that certainly match better the empirical data but whose associated jump arrival intensity is in general no more a Markov process. This is of course a serious problem, because one cannot use standard tools of stochastic calculus anymore.

the applicability of the framework to the pricing of derivatives. We conclude in section 5.

2 A new multifactor SEJD model for the equity price

2.1 The framework

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a right continuous filtration $\{\mathcal{F}_t\}_{t\geq 0}$ on which is defined the price process of a stock index denoted by $S = (S_t)_t$. We assume that the stock index price is governed by the following stochastic differential equation

$$\frac{dS_t}{S_{t^-}} = \mu dt + \sigma dW_t + d\left(\sum_{j=1}^{N_t}\left(e^{J_j}-1\right)\right) - \lambda_t \mathbb{E}\left(e^J - 1\right)dt \quad (1)$$

$$= \mu dt + \sigma dW_t + \left(e^J - 1\right)dN_t - \lambda_t \mathbb{E}\left(e^J - 1\right)dt \quad (2)$$

where μ is a constant drift term, dW_t stands for the increment of the Brownian motion $W = (W_t)_t$, σ is a constant diffusion coefficient ($\sigma \in \mathbb{R}^+$) and the final terms of the above stochastic differential equations highlight the influence of jumps that can occur in the interval $(t^-, t + dt)$. The second equality arises naturally because the probability to observe more than one jump in an infinitesimal period is negligible[9]. The last equation may also be rewritten

$$d\ln S_t = \left(\mu - \frac{\sigma^2}{2} - \lambda_t \mathbb{E}\left(e^J - 1\right)\right)dt + \sigma dW_t + J dN_t$$

The process $N = (N_t)_t$ informs on the number of jumps observed before time t, the sequence of i.i.d. random variables $(J_j)_j$ on the size of the return jumps. Each random size (J_j) is an independent copy of a random variable J. We denote by μ_J the expected jump size in absolute term, i.e. $E(|J|) = \mu_J$ which is of course a positive real number. In numerical applications, several distributions for J are tested: positive and negative exponential, double exponential and binomial laws. Double exponential jumps (DEJ)[10] are defined by three parameters (p, ρ^+, ρ^-) and their features are reminded in appendix A. Strictly positive or negative jumps are particular case of the DEJ. But testing them, allows us to emphasize later that self-excitation is not only by negative jumps as in Carr and Wu (2016) but also by positive ones. Binomial jumps, also defined by parameters (p, ρ^+, ρ^-), are constant and take either a positive or a negative value. The solution of the previous stochastic differential equation is

$$S_t = S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t - \mathbb{E}\left(e^J - 1\right)\int_0^t \lambda_s ds + \sigma W_t + \sum_{j=1}^{N_t} J_j\right)$$

[9] The first equality is useful for estimation purposes and in particular to find an appropriate Euler approximation because the time interval between two discretely sampled data is not necessarily infinitesimal.

[10] Assuming a constant jump arrival intensity for the point process N makes our model specification similar to the one introduced by Kou (2002). Actually, this author introduces the double exponential distribution to model jumps in returns that can be positive or negative. For the reader's convenience, we recall in the appendix some general results about this nowadays standard distribution. It is worth emphasizing that the Lévy process considered by Kou (2002) is unable to accomodate the clustering of jumps documented by Aït-Sahalia et al (2015) and others.

The jump arrival intensity, denoted by $\lambda = (\lambda_t)_t$, evolves over time according to the stochastic differential equation

$$d\lambda_t = \alpha(\theta_t - \lambda_t)\,dt + \eta d\left(\sum_{i=1}^{N_t} |J_i|\right) \tag{3}$$

$$= \alpha(\theta_t - \lambda_t)\,dt + \eta\,|J|\,dN_t. \tag{4}$$

The jump arrival intensity hence reverts with a reversion speed $\alpha \in \mathbb{R}^+$ to a long-run mean level θ_t that is time-varying. As jumps may be positive or negative, the variation of intensity is proportional to the absolute value of the jump of price. In Aït-Sahalia et al. (2015), the shock on intensity caused by a jump is constant and depends only on the counting process N_t. Instead, our model links this shock to the amplitude of jumps. To capture our intuition that past realized jumps in the underlying price process may impact both the instantaneous and long run propensities of observing jumps, one assumes that the behavior of the time-varying long-run mean θ_t is itself mean-reverting, self-exciting and described by

$$d\theta_t = \beta(\gamma - \theta_t)\,dt + \delta d\left(\sum_{i=1}^{N_t} |J_i|\right). \tag{5}$$

$$= \beta(\gamma - \theta_t)\,dt + \delta\,|J|\,dN_t \tag{6}$$

where $\beta \in \mathbb{R}^+$ is the speed of mean reversion and $\gamma \in \mathbb{R}^+$ the level toward which the long-run mean reverts.

The system of stochastic differential equations (1), (3) and (5) fully characterizes our multifactor self-exciting jump diffusion model. This setting deserves some general comments. First, it is clear that any jump can impact directly and indirectly the price process. Second, equations (3) and (5) show that the realized jumps in the underlying price process can simultaneously modify the jump arrival intensity $\lambda = (\lambda_t)_t$ and its long-run mean level $\theta = (\theta_t)_t$. Nevertheless, $\eta \in \mathbb{R}^+$ and $\delta \in \mathbb{R}^+$ are two constant parameters tuning the influence of past realized jumps on these two processes[11]. Third, equations (7) and (9) are both satisfied because the probability to observe more than one jump in an infinitesimal period is almost surely zero. Fourth and in case we can identify the sequence of event/jump times by $(T_n)_{n\geq 0}$, the counting process $N = (N_t)_t$ may be written by $N_t = \sum_{n\geq 1} 1_{\{T_n<t\}}$, where 1_A is the indicator function of event A. Fifth and lastly, if one considers the marked point process $L = (L_t)_t$ defined by $L_t = \sum_{i=1}^{N_t} |J_i| = \sum_{n\geq 1} |J_n|\, 1_{\{T_n<t\}}$ i.e. the sum of the absolute values of jumps in the asset price up to time t. Then, one may rewrite the jump

[11] Consequently, we can investigate their respective value and whether they are equal. If these parameters are different from zero, then any realization of a jump immediately increase the instantaneous probability of observing an other jump and the long-run level toward which this intensity process tends to mean-revert. And testing whether η and δ are different from zero is equivalent to testing whether self-excitation matters in these processes. If these parameters are equal, then any realization of a jump impact λ and θ simultaneously.

terms in equations (3) and (5) in terms of the increment of L denoted by dL_t. Moreover, a direct integration of equation (5) leads to

$$\theta_t = \gamma + (\theta_0 - \gamma)\, e^{-\beta t} + \int_0^t \delta e^{-\beta(t-s)} dL_s.$$

The exponential function ($e^{-\beta(t-s)}$) in this expression highlights that the influence of a realized jump on the long-run mean level θ decays over time exponentially. This solution must be then plugged into equation (3) and, in the next section, we can derive the analytical expression for the jump arrival intensity when the long-run mean level is stochastic. If θ_t were constant and say equal to θ, then the solution to equation (3) would simply be

$$\lambda_t = \theta + (\lambda_0 - \theta)\, e^{-\alpha t} + \int_0^t \eta e^{-\alpha(t-s)} dL_s \tag{7}$$

and the exponential function $\eta e^{-\alpha(t-u)} \equiv \phi^{\text{single}}(u)$ would be the kernel function (or memory) of the jump arrival intensity process. It must be stressed that this latter expression may seem familiar to whom knows the Hawkes process, but this process is not standard at all, Remind that the marked point process L depends on the realized jumps in *absolute* terms. Consequently, we will provide in the rest of this article results for the general process as well as for this simpler but new mono-factor process.

2.2 Analyzing the jump feature of the multifactor SEJD model

This section focuses on the jump dimension of our multifactor SEJD model. We provides various results and analytical expressions that can help understanding the main features of our general model. Before entering the core of this section, it is important to notice that the jump arrival intensity process $\lambda = (\lambda_t)_t$ is not Markov, but the joint process (λ, θ, N) is. So, to apply standard results of stochastic calculus, one must often consider the processes $\lambda = (\lambda_t)_t$, $\theta = (\theta_t)_t$ and $N = (N_t)_t$ jointly. We study first $\lambda = (\lambda_t)_t$, $\theta = (\theta_t)_t$ that are the state variables that govern the jump process. But then we consider the joint process (λ, θ, N) which is equivalent to (λ, θ, $\int \lambda_s ds$). We finally say some words of any regular function of our state variables.

Proposition 1 provides an analytical expression for the value of the jump arrival intensity at time t in our multifactor self-exciting jump diffusion model. It reveals among other things that the kernel function of the jump arrival intensity is the sum of two exponential functions.

Proposition 1: *The jump arrival intensity of the multifactor self-excited process admits the following expression*

$$\lambda_t = \gamma + (\lambda_0 - \gamma) e^{-\alpha t} + (\theta_0 - \gamma) \frac{\alpha}{\alpha - \beta} \left(e^{-\beta t} - e^{-\alpha t} \right) + \int_0^t \phi^{\text{bi}}(u)\, dL_u \tag{8}$$

where the kernel function is

$$\phi^{\text{bi}}(u) = \frac{\alpha\delta}{\alpha - \beta} e^{-\beta(t-u)} - \left(\frac{\alpha\delta}{\alpha - \beta} - \eta \right) e^{-\alpha(t-u)}$$

if $\alpha \neq \beta$ *and*

$$\lambda_t = \gamma + (\lambda_0 - \gamma) e^{-\alpha t} + (\theta_0 - \gamma) \alpha t e^{-\alpha t} + \int_0^t \phi(u)\, dL_u$$

$$\phi(u) = \delta\alpha e^{-\alpha(t-u)} (t-u).$$

otherwise.

Proposition 1 highlights that the kernel function of the jump arrival intensity, in our multifactor setting, is a (weighted) sum of two exponential functions. This kernel function of the multifactor jump process (the multifactor kernel function for short) may also be rewritten

$$\begin{aligned} \phi^{\text{bi}}(u) &= \eta e^{-\alpha(t-u)} + \delta \frac{\alpha}{\alpha - \beta} \left(e^{-\beta(t-u)} - e^{-\alpha(t-u)} \right) \\ &= \eta e^{-\alpha(t-u)} + \delta e^{-\beta(t-u)} \alpha \left[\frac{1 - e^{-(\alpha-\beta)(t-u)}}{\alpha - \beta} \right], \text{ for } u < t \end{aligned}$$

The last expression highlights the meaningful functions $\eta e^{-\alpha(t-u)}$ and $\delta e^{-\beta(t-u)}$. These expressions in turn calls for two remarks. First, the multifactor kernel function nests the function $\eta e^{-\alpha(t-u)}$ as a special case (just set δ to zero), meaning that the multifactor jump process can capture a larger spectrum of stochastic behaviors than the single factor jump process. Second, the second term of the r.h.s. $(\delta \frac{\alpha}{\alpha-\beta}\left(e^{-\beta(t-u)} - e^{-\alpha(t-u)}\right))$ is positive whatever the values of α and β. So, assuming that all parameter estimates can be similar, the multifactor kernel function will be at least as large as the single factor one, meaning that the influence of a realized jump on the intensity may last longer in our bifactor setting. Of course, this cannot be the case because the parameter estimates will compensate for this effect. Actually, both models tend to capture the same reality... so that we can anticipate that in the empirical investigations, the single factor kernel function tend to "capture', summarize' or 'average' the information content of the multifactor kernel function.

Proposition 2: *If* $(\beta - (\eta\mu_J - \alpha))^2 + 4(\delta\mu_J\alpha + \beta(\eta\mu_J - \alpha)) \geq 0$*, the expectations of* θ_t *and* λ_t *are given by*

$$\begin{pmatrix} \mathbb{E}(\theta_t|\mathcal{F}_0) \\ \mathbb{E}(\lambda_t|\mathcal{F}_0) \end{pmatrix} = V \begin{pmatrix} \frac{1}{\gamma_1}\left(e^{\gamma_1 t} - 1\right) & 0 \\ 0 & \frac{1}{\gamma_2}\left(e^{\gamma_2 t} - 1\right) \end{pmatrix} V^{-1} \begin{pmatrix} \gamma\beta \\ 0 \end{pmatrix} + V \begin{pmatrix} e^{\gamma_1 t} & 0 \\ 0 & e^{\gamma_2 t} \end{pmatrix} V^{-1} \begin{pmatrix} \theta_0 \\ \lambda_0 \end{pmatrix} \tag{9}$$

where γ_1*,* γ_2 *are constant given by*

$$\begin{aligned} \gamma_1 &= \frac{1}{2}\left((\eta\mu_J - \alpha) - \beta\right) + \frac{1}{2}\sqrt{(\beta - (\eta\mu_J - \alpha))^2 + 4(\delta\mu_J\alpha + \beta(\eta\mu_J - \alpha))} \\ \gamma_2 &= \frac{1}{2}\left((\eta\mu_J - \alpha) - \beta\right) - \frac{1}{2}\sqrt{(\beta - (\eta\mu_J - \alpha))^2 + 4(\delta\mu_J\alpha + \beta(\eta\mu_J - \alpha))} \end{aligned} \tag{10}$$

and V *is a matrix defined by*

$$V = \begin{pmatrix} -\delta\mu_J & -\delta\mu_J \\ -\beta - \gamma_1 & -\beta - \gamma_2 \end{pmatrix}$$

whose determinant and inverse are respectively equal to $\Upsilon = \delta\mu_J(\gamma_2 - \gamma_1)$ *and* $V^{-1} = \frac{1}{\Upsilon}\begin{pmatrix} -\beta - \gamma_2 & \delta\mu_J \\ \beta + \gamma_1 & -\delta\mu_J \end{pmatrix}$. An important consequence of this proposition is that the model is stable (in the sense that the limits of λ_t and θ_t exist when $t \to +\infty$) if and only if γ_1 and γ_2 are negative. In view of expressions (10), this means that the sum $(\alpha + \beta)$ is significantly larger than $\eta\mu_J$. Hereafter, we constrain structural parameters so as to imply negative values for γ_1 and γ_2 and to ensure a stable jump model. We will see later that α is constrained to be larger than $\eta\mu_J$. Throughout the rest of the paper, we assume that

$\alpha \neq \beta$. Under these conditions, the above asymptotic expressions become

$$\lim_{t\to\infty} \begin{pmatrix} \mathbb{E}(\theta_t|\mathcal{F}_0) \\ \mathbb{E}(\lambda_t|\mathcal{F}_0) \end{pmatrix} = V \begin{pmatrix} -\frac{1}{\gamma_1} & 0 \\ 0 & -\frac{1}{\gamma_2} \end{pmatrix} V^{-1} \begin{pmatrix} \gamma\beta \\ 0 \end{pmatrix} \tag{11}$$

$$= \begin{pmatrix} \gamma\beta\left(-\beta\frac{1}{\gamma_1\gamma_2} - \frac{(\gamma_2+\gamma_1)}{\gamma_1\gamma_2}\right) \\ -(\beta+\gamma_2)(\beta+\gamma_1)\frac{\gamma\beta}{\delta\mu_J}\left(\frac{1}{\gamma_1\gamma_2}\right) \end{pmatrix}$$

where γ_1 and γ_2 are defined in proposition 2.2. Plugging there $(\beta+\gamma_2)(\beta+\gamma_1) = -\delta\mu_J\alpha$ and $\gamma_1\gamma_2 = \beta(\alpha-\eta\mu_J) - \delta\mu_J\alpha$ (computed in the appendix) allows us to rewrite

$$\lim_{t\to\infty} \begin{pmatrix} \mathbb{E}(\theta_t|\mathcal{F}_0) \\ \mathbb{E}(\lambda_t|\mathcal{F}_0) \end{pmatrix} = \begin{pmatrix} \gamma\frac{\beta(\alpha-\eta\mu_J)}{\beta(\alpha-\eta\mu_J)-\alpha\delta\mu_J} \\ \gamma\frac{\alpha\beta}{\beta(\alpha-\eta\mu_J)-\alpha\delta\mu_J} \end{pmatrix}.$$

And a couple of comments deserve to be done. First, because we constrain γ_1 and γ_2 to be negative, $\gamma_1\gamma_2 = \beta(\alpha-\eta\mu_J) - \delta\mu_J\alpha$ is positive and $(\alpha-\eta\mu_J)$ too. So these asymptotic values are well-defined. In addition, they share a common term so that we can write

$$\lim_{t\to\infty}(\mathbb{E}(\theta_t|\mathcal{F}_0)) = \lim_{t\to\infty}\mathbb{E}(\lambda_t|\mathcal{F}_0) - \frac{\beta\eta\mu_J}{\beta(\alpha-\eta\mu_J)-\alpha\delta\mu_J}$$

and, by design, the asymptotic value for θ is lower than the one for λ. Finally, the equation (9) may be rewritten

$$\begin{aligned} \mathbb{E}(\lambda_t|\mathcal{F}_0) &= \frac{1}{\Upsilon}\left[(\beta+\gamma_2)(\beta+\gamma_1)\left(\frac{\gamma\beta}{\gamma_1}+\theta_0\right) - (\beta+\gamma_1)\delta\mu_J\lambda_0\right]e^{\gamma_1 t} - \\ &\quad \frac{1}{\Upsilon}\left[(\beta+\gamma_2)(\beta+\gamma_1)\left(\frac{\gamma\beta}{\gamma_2}+\theta_0\right) - (\beta+\gamma_2)\delta\mu_J\lambda_0\right]e^{\gamma_2 t} + \\ &\quad \frac{1}{\Upsilon}(\beta+\gamma_2)(\beta+\gamma_1)\left(\frac{\gamma\beta}{\gamma_2} - \frac{\gamma\beta}{\gamma_1}\right) \\ : &= \upsilon_1 e^{\gamma_1 t} - \upsilon_2 e^{\gamma_2 t} + \upsilon_3. \end{aligned} \tag{12}$$

to highlight that the expected value of λ_t is an affine function of exponentials.

The next proposition provides a general decomposition (in two terms) of the variance of the jump arrival intensity and it provides an explicit expression for it. This decomposition highlights the key role of the kernel function in the computation of the variance. So we expect the variance to be significantly influenced by the shape of the kernel function.

Proposition 3: *The variance of the jump arrival intensity λ_t may be written*

$$\mathbb{V}(\lambda_t|\mathcal{F}_0) = \mathbb{E}(J^2)\int_0^t \phi^{\text{biv}}(u)^2\,\mathbb{E}(\lambda_u|\mathcal{F}_0)\,du$$

where

$$\int_0^t \phi^{\text{bi}}(u)^2 \, \mathbb{E}(\lambda_u|\mathcal{F}_0)\, du = \varepsilon_3 \frac{\upsilon_3}{\alpha+\beta}\left(1-e^{-(\alpha+\beta)t}\right) + \tag{13}$$
$$\varepsilon_1\left[\frac{\upsilon_1}{\gamma_1+2\beta}\left(e^{\gamma_1 t}-e^{-2\beta t}\right)-\frac{\upsilon_2}{\gamma_2+2\beta}\left(e^{\gamma_2 t}-e^{-2\beta t}\right)+\frac{\upsilon_3}{2\beta}\left(1-e^{-2\beta t}\right)\right]+$$
$$\varepsilon_2\left[\frac{\upsilon_1}{\gamma_1+2\alpha}\left(e^{\gamma_1 t}-e^{-2\alpha t}\right)-\frac{\upsilon_2}{\gamma_2+2\alpha}\left(e^{\gamma_2 t}-e^{-2\alpha t}\right)+\frac{\upsilon_3}{2\alpha}\left(1-e^{-2\alpha t}\right)\right]+$$
$$\varepsilon_3\left[\frac{\upsilon_1}{\gamma_1+\alpha+\beta}\left(e^{\gamma_1 t}-e^{-(\alpha+\beta)t}\right)-\frac{\upsilon_2}{\gamma_2+\alpha+\beta}\left(e^{\gamma_2 t}-e^{-(\alpha+\beta)t}\right)\right]$$

with $\varepsilon_1 = \left(\frac{\alpha\delta}{\alpha-\beta}\right)^2$, $\varepsilon_2 = \left(\eta - \frac{\alpha\delta}{\alpha-\beta}\right)^2$ *and* $\varepsilon_3 = 2\left(\frac{\alpha\delta}{\alpha-\beta}\right)\left(\eta-\frac{\alpha\delta}{\alpha-\beta}\right)$. *When the jump size has a binomial, simple or double exponential distribution, its second moment* $\mathbb{E}(J^2)$ *admits a closed form (see appendix A).*

Whether the process is stable γ_1 and γ_2 are negative, the limit of $\mathbb{V}(\lambda_t|\mathcal{F}_0)$ for $t \to \infty$ is constant as all exponential terms in equation (13) decay to zero:

Corollary 1: *The asymptotic variance of* λ_t *is independent from the initial value for the process and equal to*

$$\lim_{t\to\infty} \mathbb{V}(\lambda_t|\mathcal{F}_0) = -\mathbb{E}(J^2)\,\frac{\gamma\beta(\beta+\gamma_2)(\beta+\gamma_1)}{\delta\mu_J\gamma_2\gamma_1}\times$$
$$\left[2\left(\frac{\alpha\delta}{\alpha^2-\beta^2}\right)\left(\eta-\frac{\alpha\delta}{\alpha-\beta}\right)+\left(\frac{\alpha\delta}{\alpha-\beta}\right)^2\frac{1}{2\beta}+\left(\eta-\frac{\alpha\delta}{\alpha-\beta}\right)^2\frac{1}{2\alpha}\right]$$

The probability of observing no jump over an interval of time $[0,T]$ may be computed by considering the following expression

$$\mathbb{P}[N_T=0] = \mathbb{E}\left[\exp\left(-\int_0^t \lambda_u du\right)|\mathcal{F}_0\right].$$

Actually, this expression is just a special case of the joint moment generating function of the triplet $\left(\Lambda_t \equiv \int_0^t \lambda_u du, \lambda_t, \theta_t\right)$. The next proposition provides a semi closed form expression for this useful function.

Proposition 4: *Let us define* $\psi(z_1,z_2) := \mathbb{E}\left(e^{z_1 J + z_2|J|}\right)$. *The joint moment generating function of* Λ_s, λ_s *and* θ_s *is given by*

$$\mathbb{E}\left(e^{\omega_0\Lambda_s+\omega_1\lambda_s+\omega_2\theta_s}\mid\mathcal{F}_t\right) = \exp\left(A(t,s)+B(t,s)\lambda_t+C(t,s)\theta_t+\omega_0\Lambda_t\right).$$

where $A(t,s)$, $B(t,s)$ and $C(t,s)$ satisfy the system of ODE's

$$\begin{cases} \frac{\partial A}{\partial t} & = -\beta\gamma C \\ \frac{\partial B}{\partial t} & = \alpha B - \left[\psi\left(0, B\eta + C\delta\right) + \omega_0 - 1\right] \\ \frac{\partial C}{\partial t} & = -\alpha B + \beta C \end{cases} . \tag{14}$$

with the terminal conditions $A(s,s) = 0$, $B(s,s) = \omega_1$ and $C(s,s) = \omega_2$.

The closed form expression of $\psi(z_1, z_2)$ is provided in appendix. There does not exist any closed form expressions for $A(t,s)$, $B(t,s)$ and $C(t,s)$. However, one may compute them by an Euler's method. The result of Proposition 4 is especially useful because one can now compute a number of by-products such as the probability of no jump over a certain period of time (see above), the probability density function of λ_t or of θ_t, etc. Notice also that proposition 4 serves us in a following section to identify the dynamics of jumps under affine equivalent measures. Before closing this passage on the state variables, it is worth mentioning some available results for any stochastic process $Y = (Y_t)_t$ defined by $Y_t = f(t, \lambda_t, \theta_t)$ where f is a regular function of time, of the intensity and of the mean reversion level. Firstly, the infinitesimal generator associated to the underlying jump process is

$$\begin{aligned} \mathcal{A}f &= f_t + \alpha(\theta_t - \lambda_t)\, f_\lambda + \beta\left(\gamma - \theta_t\right) f_\theta \\ &\quad + \lambda_t \int_0^{+\infty} \left[f\left(t,\, \lambda_t + \eta\,|z|, \theta_t + \delta\,|z|\right) - f\left(t,\, \lambda_t, \theta_t\right)\right]\, d\nu(z). \end{aligned}$$

Second, the dynamics of Y can be described by

$$dY_t = \left[f_t + \alpha(\theta_t - \lambda_t)\, f_\lambda + \beta\left(\gamma - \theta_t\right) f_\theta\right] dt + \left[f\left(t,\, \lambda_t + \eta\,|J|, \theta_t + \delta\,|J|\right) - f\left(t,\, \lambda_t, \theta_t\right)\right] dN_t.$$

We can have a look at the equity price. Next proposition first of derives derives the moment generating function of the log return of S_t denoted by $X_t := \ln \frac{S_t}{S_0}$.

Proposition 5: *If $\psi(z_1, z_2) := \mathbb{E}\left(e^{z_1 J + z_2 |J|}\right)$, the moment generating function of $\omega_1 X_s$ for $s \geq t$, is given by*

$$\mathbb{E}\left(e^{\omega_1 X_s} \,|\, \mathcal{F}_t\right) = \left(\frac{S_t}{S_0}\right)^{\omega_1} \exp\left(A(t,s) + B(t,s)\lambda_t + C(t,s)\theta_t\right).$$

where $A(t,s)$, $B(t,s)$ and $C(t,s)$ are solutions of the system of ODE's

$$\begin{cases} \frac{\partial A}{\partial t} & = -\omega_1\left(\mu - \frac{\sigma^2}{2}\right) - \omega_1^2 \frac{\sigma^2}{2} - \beta\gamma C \\ \frac{\partial B}{\partial t} & = \alpha B + \omega_1\left(\psi\left(1, 0\right) - 1\right) - \left[\psi\left(\omega_1\,,\, B\,\eta + C\,\delta\right) - 1\right] \\ \frac{\partial C}{\partial t} & = -\alpha B + \beta C \end{cases} \cdot \tag{15}$$

with the terminal conditions $A(s,s) = 0$, $B(s,s) = 0$ and $C(s,s) = 0$.

The next proposition provides an explicit expression for the realized variance and its expectation as measured by the quadratic variation of instantaneous returns.

Proposition 6: *The realized variance, measured as the quadratic variation of* $\int_0^t \frac{dS_s}{S_s}$ *over the period* $[0, T]$ *is given by*

$$\sigma_R^2(T) = \sigma^2 T + \int_0^T \left(e^{J_u} - 1\right)^2 dN_u. \tag{16}$$

Its expectation is equal to

$$\mathbb{E}\left(\sigma_R^2(T) \,|\, \mathcal{F}_0\right) = \sigma^2 T + \mathbb{E}\left(\left(e^J - 1\right)^2\right) \int_0^T \mathbb{E}\left(\lambda_u \,|\, \mathcal{F}_0\right) du \tag{17}$$

where $\int_0^T \mathbb{E}\left(\lambda_u \,|\, \mathcal{F}_0\right) du$ *is provided by equation (12).* $\mathbb{E}\left(\left(e^J - 1\right)^2\right)$ *is provided in appendix A for binomial, simple and double exponential jumps.*

This last proposition reveals that the curve of expected realized variances is a mixture of exponential functions. The expected realized volatility does not admit any semi-closed form expression. However Brockhaus and Long (2000) suggest the following approximation, that is the second order expansion for $\sqrt{x}$, for its evaluation:

$$\mathbb{E}\left(\sqrt{\sigma_R^2(T)} \,|\, \mathcal{F}_0\right) = \sqrt{\mathbb{E}\left(\sigma_R^2(T) \,|\, \mathcal{F}_0\right)} - \frac{\mathbb{V}\left(\sigma_R^2(T) \,|\, \mathcal{F}_0\right)}{8\,\mathbb{E}\left(\sigma_R^2(T) \,|\, \mathcal{F}_0\right)^{\frac{3}{2}}}$$

where the variance of the realized variance is provided by the next proposition

Proposition 7: *The variance of* $\sigma_R^2(T)$ *is equal to*

$$\mathbb{V}\left(\sigma_R^2(T) \,|\, \mathcal{F}_0\right) = \frac{1}{T^2}\mathbb{E}\left(\left(e^J - 1\right)^4\right) \int_0^T \mathbb{E}\left(\lambda_u \,|\, \mathcal{F}_0\right) du$$

where $\int_0^T \mathbb{E}\left(\lambda_u \,|\, \mathcal{F}_0\right) du$ *is provided by equation (12).* $\mathbb{E}\left(\left(e^J - 1\right)^4\right)$ *is provided in appendix A for binomial, simple and double exponential jumps.* This result is a direct consequence of the expression (16) for the realized variance.

2.3 The single factor model

It is now interesting to stress the difference of our multifactor self-exciting jump multifactor model with the model where θ_t is a constant θ. Let's have a look at the different expressions. The expected jump arrival intensity is then given by

$$\mathbb{E}\left(\lambda_t \,|\, \mathcal{F}_0\right) = \lambda_0 e^{-(\alpha - \eta\mu_J)t} + \lambda_\infty \left(1 - e^{-(\alpha - \eta\mu_J)t}\right) \tag{18}$$

where $\lambda_\infty = \frac{\alpha}{\alpha - \eta\mu_J}\theta$. This expression is well defined and meaningful if $\alpha - \eta\mu_J > 0$. As expected, the expected value of the jump arrival intensity tends to a constant as t becomes large. Some readers may be surprised that this long-run mean level is $\lambda_\infty = \frac{\alpha}{\alpha - \eta\mu_J}\theta$ and not θ. The reason for this is that the jump term in the dynamics of λ is not compensated and it thus has no zero expectation. The equivalent compensated process would have the following dynamics:

$$d\lambda_t = a\left(\lambda_\infty - \lambda_t\right) dt + \left(\eta\left|J\right| dN_t - \lambda_t \eta\mu_J dt\right)$$

but may lead to a negative intensity. This expression is far simpler than the one associated to the multifactor jump process. The variance of the jump arrival intensity satisfies

$$\mathbb{V}\left(\lambda_t|\mathcal{F}_0\right) = \mathbb{E}(J^2)\int_0^t \phi^{\text{single}}\left(u\right)^2 \mathbb{E}\left(\lambda_u|\mathcal{F}_0\right) du$$

where $\mathbb{E}\left(\lambda_u|\mathcal{F}_0\right)$ is given by equation (18) and $\phi^{\text{single}}\left(u\right) = \eta e^{-\alpha(t-u)}$. This expression is similar to the one provided by Proposition 3. And after simplification, the variance has the following form

$$\mathbb{V}\left(\lambda_t|\mathcal{F}_0\right) = \mathbb{E}(J^2)\left[\frac{\eta^2(1-\lambda_\infty)}{\alpha+\eta\mu_J}\left(e^{-(\alpha-\eta\mu_J)t} - e^{-2\alpha t}\right) + \frac{\eta^2\lambda_\infty}{2\alpha}\left(1 - e^{-2\alpha t}\right)\right],$$

that converges asymptotically to the next constant:

$$\lim_{t\to\infty}\mathbb{V}\left(\lambda_t|\mathcal{F}_0\right) = \mathbb{E}(J^2)\frac{\eta^2\lambda_\infty}{2\alpha}.$$

When the mean reversion level of λ_t is constant and equal to θ, the moment generating function of the log return,X_t is given by

$$\mathbb{E}\left(e^{\omega_1 X_s}\,|\,\mathcal{F}_t\right) = \left(\frac{S_t}{S_0}\right)^{\omega_1}\exp\left(A(t,s) + B(t,s)\lambda_t\right),\ s \geq t$$

where $A(t,s)$ and $B(t,s)$ are solutions of the system of ODE's

$$\begin{cases} \frac{\partial}{\partial t}A &= -\omega_1\left(\mu - \frac{\sigma^2}{2}\right) - \omega_1^2\frac{\sigma^2}{2} - \alpha\theta B \\ \frac{\partial}{\partial t}B &= \alpha B + \omega_1\left(\psi\left(1,0\right) - 1\right) - \left[\psi\left(\omega_1\,,\,B\,\eta\right) - 1\right] \end{cases}. \tag{19}$$

with the terminal conditions $A(s,s) = 0$, $B(s,s) = 0$. Finally, the expression of the realized variance is similar to (17), but in which the integral of the expected intensity is given by:

$$\int_0^T \mathbb{E}\left(\lambda_u\,|\,\mathcal{F}_0\right) du = \frac{\lambda_0 - \lambda_\infty}{\alpha - \eta\mu_J}\left(1 - e^{-(\alpha-\eta\mu_J)T}\right) + \lambda_\infty T$$

.

3 Estimating multifactor SEJD.

In this section, we apply our two SEJD models to financial data. We use only returns data to estimate and test the models. Our main goal is to test in-sample whether and how considering a self-exciting process for the long-run mean level of the jump arrival intensity can help matching the dynamics of data. The other purpose of this analysis is to select the most appropriate distributions for jumps sizes, among the positive, negative, double exponential and binomial laws. Before presenting some empirical results, we must describe our data and introduce the econometric strategy we develop to filter state variables and to estimate the parameters. Notice that we only provide hereafter a sketch of our econometric approach for estimating structural parameters and for filtering hidden intensities. These approaches are fully described in the Appendix devoted to the econometric methodology.

3.1 Data Description

We collect S&P 500 daily data from Bloomberg over a sample period from September 2005 to October 2015. As a result, the time serie contains a total of 2543 continuously compounded returns. Table 1 provides summary statistics for the continuously compounded returns. Their yearly volatility reaches 20.64% and the very high kurtosis indicates that the distribution of returns has fat-tails. Jarque Bera and Lillie tests of normality reject this assumption whereas the Durbin Watson statistic reveals serial dependence.

Insert Table 1

Figure 1 plots prices and returns of the 'assets' on the sample period. Clustering of jumps in returns are clearly visible from September 2008 to end 2009 (the US credit crunch period) and from September 2011 to February 2012 (the second period of the double-dip recession). Shocks on returns during these periods do not display any clear trend: negative movements alternate regularly with large positive technical bounces. This observation corroborates a link between the frequency of jumps and their absolute values, as assumed in our modelling.

Insert Figure 1

3.2 Structural parameter estimation

It is common knowledge that estimating parameters of a Jump-Diffusion process with time series is challenging and requires some advanced econometric techniques. Our model involves, in particular, two latent processes ($\lambda = (\lambda_t)_t$ and $\theta = (\theta_t)_t$) that we need to back out by using a filtering technique. An Euler discretization makes it possible to rewrite our model into a state space form[12]. The state-space formulation we obtain is nevertheless highly non-Gaussian and nonlinear, so that we cannot use neither the popular Kalman filter nor any variants in the present context. There is a considerable literature on how to perform filtering of non-Gaussian and nonlinear state space models. In this paper we employ a peak over threshold approach. This approach detailed in appendix is particularly robust and simple to implement. It is also enough accurate to emphasize the main advantages of the bifactor model compared to the single factor one.

After calibration of parameters, a sequential Monte Carlo procedure (also called particles filter), running with 5000 particles, is applied so as to approach the loglikelihood and to filter with more accuracy the state variables. Notice that the loglikelihood function yield by this algorithm is obtained by simulations and is then not smooth enough to calibrate directly the model by loglikelihood maximization. A way to avoid this drawback consists to use a Particle Markov Chain Monte Carlo (PMCMC) method. But to limit the computation time, this approach requires to determine accurately prior and transition distributions of parameters, that we don't know.

3.3 Empirical results

Table 2 reports estimated log-likelihoods, Akaike Information Criterions (AIC) for the uni- and bi-factor models, with simple or double exponential and binomial jumps. Comparing AIC statistics suggests that the bifactor model outperforms the single factor one to fit the data, excepted whether jumps have positive exponential distributions. The same conclusion arises by observing larger log-likelihoods for the multifactor model. But are these likelihoods significantly larger? To answer this question, it is tempting to 'test' whether this result is true. Notice however these are just average log-likelihoods so that the test is only approximate. In the final row of Table 2, we routinely apply the likelihood ratio test to assess whether the multifactor model really outperforms the single factor model. The likelihood ratio statistic is commonly defined by $LRT = -2\left(\ln L_0 - \ln L_a\right)$ where $\ln L_0$ is the log-likelihood of the single factor specification and $\ln L_a$ is the log-likelihood of the multifactor specification. This test statistic behaves asymptotically like a random variable having a chi-squared distribution whose degree of freedom is d where d is the difference of the number of parameters involved in the two specification. One therefore has $LRT \to \chi^2_d$ where $d = 2$ This analysis leads to the same conclusion. The best fit is obtained with DEJ jumps. This suggests that self-excitation is caused by positive and negative jumps and not only by negative shocks as in Carr and Wu (2016). The asymptotic frequency is around 22 jumps per year for models with double exponential jumps.

Insert Table 2

[12]This discretization can of course induce a bias. But this bias is expected to be quite small because we apply our model to daily data.

Table 3 reports parameters for the uni- and bi-factor specifications. Positive shocks occur with a probability of 37% and their average amplitude is 3.28%. Negative jumps are less frequent with a slightly larger average amplitude (-2.95%). By construction, the parameter η tunes the feedback effect of jumps on λ_t. The η of bifactor models are all below the one of their equivalent in the unifactor model. Speeds of mean reversion of λ_t are higher in bifactor models then in single factor ones. Levels of mean reversion γ and θ are quite similar. Table 4 presents the asymptotic expectation and standard deviations of λ_t and θ_t, for the uni- and bi-factor models.

Insert Table 3

Insert Table 4

Figure 2 provides the QQ plots to assess the quality of the DEJ multifactor model to fit the data. The left upper graph is a standard normality plot for assessing whether residuals are Gaussian. The other graph questions whether the filtered jumps behave according to a double exponential distribution.

Insert Figure 2

Insert Figure 3

The figure 3 compares the kernel functions of bi- and uni-variate models with double constant and double exponential jumps. The graph in the second and third rows show respectively the spread between double and single factor kernels and their ratio. The kernel function at time zero informs on the instantaneous impact of a price jump on the jump arrival intensity. And graphs reveal that this value is much higher for the bifactor model ($\eta = 381$) than for the single factor one ($\eta = 337$). As time passes a bit, the multifactor kernel function decreases faster than the single factor one. After 0.08 year, the multifactor kernel function decreases slower than the single factor one, so that the influence of a realized jump in the stock index or the individual index lasts longer in the multifactor model. This is confirmed by ratios double on single factor kernels that emphasizes the fatter left tails of double factor kernel.

Insert Figure 4

The figure 4 shows the term structure of expected realized volatilities yield by the bi-factor model. We observe that the closer is λ_0 to the asymptotic level $\mathbb{E}(\lambda_\infty)$, the flater is the curve. As we could expect, deviations of λ_0 mainly modify the short term slope of the curve. Deviations of θ_0 from $\mathbb{E}(\lambda_\infty)$ have a bigger impact and affect on the long run the term structure. For values of θ_0 above (resp. below) $\mathbb{E}(\theta_\infty)$, the curve is convex decreasing (resp. concave increasing).

Insert Figure 5

Thanks to our econometric methodology, we can filter out values of λ_t and θ_t over time. We present these time series in Figure 5. Because both λ_t and θ_t react to the same jumps, their evolution over time seems a bit similar. They both experiment some peaks around 2008 the period where the US experiments a credit crunch and between September 2011 and February 2012 which is the second period of the double-dip recession. But, because their relative sensitivity to the realized jumps is of different order as shown by the estimates of η vs δ in Table 3, we can rationalize some differences. In particular, the θ_t is less affected by jumps than λ_t.

4 Pricing with a SEDJ model

4.1 Changes of measure

Empirical evidences suggest that the bi-factor SEDJ is efficient for modelling of time series. This section shows that it is also usable for valuation purposes. As prices depend on two hidden state processes, a market made up of cash and of one stock is incomplete by construction. Hence, several equivalent measures are eligible as a candidate for the definition of a risk neutral one. In this paper, we focus on a family of changes of measure that preserves the dynamics of the process. These are induced by exponential martingales of the form:

$$\begin{aligned} M_t(\xi,\varphi) \quad := \quad & \exp\left((\kappa_1(\xi),\, \kappa_2(\xi)) \begin{pmatrix} \lambda_t \\ \theta_t \end{pmatrix} + \xi\, L_t - \kappa_3(\xi)t \right) \\ & \times \exp\left(-\frac{1}{2}\int_0^t \varphi(s)^2 ds - \int_0^t \varphi(s) dW_s \right) \end{aligned} \tag{20}$$

where $\varphi(s)$ is $\mathcal{F}_s$ adapted and ξ is constant. $\kappa_1(\xi)$, $\kappa_2(\xi)$ and $\kappa_3(\xi)$ are functions of ξ. Zhang et al. (2009) use a similar change of measure to simulate rare events, of a one dimension Hawkes process, without Brownian component and only with constant jumps. In our framework, jumps are random and the affine change of measure modifies both frequencies and distribution. The next proposition details the conditions under which M_t is a martingale:

Proposition 8: *If for any parameter ξ, there exist suitable solutions $\kappa_1(.)$, $\kappa_2(.)$ and $\kappa_3(.)$ for the system of equations*

$$\begin{cases} \kappa_1\alpha - \kappa_2\beta & = 0 \\ \kappa_2\beta\gamma - \kappa_3 & = 0 \\ \kappa_1\alpha - (\psi(0, \kappa_1\eta + \kappa_2\delta + \xi) - 1) & = 0 \end{cases} \tag{21}$$

then $M_t(\xi)$ is a local martingale.

Assuming the existence of suitable solutions for the system (21), an equivalent measure $Q^{\xi,\varphi}$ is defined by:

$$\left.\frac{dQ^{\xi,\varphi}}{dP}\right|_{\mathcal{F}_t} = \frac{M_t(\xi,\varphi)}{M_0(\xi,\varphi)} \tag{22}$$

The next proposition develops the dynamics of λ_t and θ_t under Q.

Proposition 9: *Let us denote by N_t^Q, the counting process that is ruled by the following intensity*

$$\lambda_t^Q = \psi(0, \kappa_1\eta + \kappa_2\delta + \xi)\lambda_t$$

under $Q^{\xi,\varphi}$. We also define random variables J^Q through the next moment generating function:

$$\psi^Q(z_1, z_2) := \mathbb{E}^Q\left(e^{z_1 J^Q + z_2|J^Q|}\right) = \frac{\psi(z_1, z_2 + (\kappa_1\eta + \kappa_2\delta + \xi))}{\psi(0, \kappa_1\eta + \kappa_2\delta + \xi)}$$

and the process $L_t^Q = \sum_{j=1}^{N_t^Q} |J_j^Q|$. Then the dynamics of λ_t and θ_t under Q^ξ is the following

$$d\lambda_t = \alpha\left(\theta_t^Q - \lambda_t\right)dt + \eta^Q dL_t^Q \tag{23}$$

$$d\theta_t = \beta\left(\gamma^Q - \theta_t\right)dt + \delta^Q dL_t^Q \tag{24}$$

where

$$\begin{aligned} \gamma^Q &= \gamma\psi(0,\, \kappa_1\eta + \kappa_2\delta + \xi) \\ \eta^Q &= \eta\psi(0,\, \kappa_1\eta + \kappa_2\delta + \xi) \\ \delta^Q &= \delta\psi(0,\, \kappa_1\eta + \kappa_2\delta + \xi) \end{aligned}$$

The next proposition shows that the jump distribution is preserved under the chosen equivalent measure.

Proposition 10: *Under $Q^{\xi,\varphi}$, jumps, J_i^Q are double-exponential random variables with density*

$$\nu^Q(z) = p^Q\rho^{+Q}e^{-\rho^{+Q}z}1_{\{z\geq 0\}} - (1-p^Q)\rho^{-Q}e^{-\rho^{-Q}z}1_{\{z<0\}},$$

where the parameters are adjusted as follows:

$$\begin{aligned} \rho^{+Q} &= \rho^{+} - (\kappa_1\eta + \kappa_2\delta + \xi)\,, \\ \rho^{-Q} &= \rho^{-} + (\kappa_1\eta + \kappa_2\delta + \xi)\,, \\ p^{Q} &= \frac{p\rho^{+}\rho^{-Q}}{(p\rho^{+}\rho^{-Q} + (1-p)\rho^{-}\rho^{+Q})}. \end{aligned}$$

As we have now identified a class of equivalent measures, we can define at least one risk neutral measure under which the expected return on S_t is equal to the risk free rate. As the dynamics of X_t under Q is driven by the following SDE

$$\begin{aligned} dX_t &= \left(\mu - \frac{\sigma^2}{2} - \lambda_t \mathbb{E}\left(e^J - 1\right) - \varphi(t)\sigma\right) dt + \sigma dW_t^Q + J^Q dN_t^Q \\ &= \left(\mu - \frac{\sigma^2}{2} - \lambda_t^Q \frac{\mathbb{E}\left(e^J - 1\right)}{\psi(0, \kappa_1\eta + \kappa_2\delta + \xi)} - \varphi(t)\sigma\right) dt + \sigma dW_t^Q + J^Q dN_t^Q \end{aligned}$$

and as $dS_t = d(e^{X_t})$, then we infer that

$$\begin{aligned} dS_t = {} & (\mu - \varphi(t)\sigma)\, S_t dt + \sigma\, S_t\, dW_t^Q \\ & + S_t \left[\left(e^{J_t^Q} - 1\right) dN_t^Q - \lambda_t^Q \frac{\mathbb{E}\left(e^J - 1\right)}{\psi(0, \kappa_1\eta + \kappa_2\delta + \xi)} dt\right] \end{aligned}$$

and

$$\mathbb{E}^Q\left(\frac{dS_t}{S_t} \,|\, \mathcal{F}_t\right) = \mu - \varphi(t)\sigma + \lambda_t^Q \left[\mathbb{E}^Q\left(e^{J^Q} - 1\right) - \frac{\mathbb{E}\left(e^J - 1\right)}{\psi(0, \kappa_1\eta + \kappa_2\delta + \xi)}\right] dt$$

This drift is equal to the risk free rate if and only if

$$\varphi(t) = \frac{\mu + \lambda_t^Q \left[\mathbb{E}^Q\left(e^{J^Q} - 1\right) - \frac{\mathbb{E}\left(e^J - 1\right)}{\psi(0, \kappa_1\eta + \kappa_2\delta + \xi)}\right] - r}{\sigma}$$

and we can summarize our result in the next corollary.

Corollary 2: *If the risk free rate, noted* r *, is constant, then the* F_t *adapted process* $\varphi(t)$ *is*

$$\varphi(t) = \frac{\mu + \lambda_t^Q \left[\mathbb{E}^Q\left(e^{J^Q} - 1\right) - \frac{\mathbb{E}\left(e^J - 1\right)}{\psi(0, \kappa_1\eta + \kappa_2\delta + \xi)}\right] - r}{\sigma}$$

to ensure the absence of arbitrage. And S_t *is the following exponential under* Q*:*

$$\begin{aligned} dS_t &= rS_t dt + \sigma\, S_t\, dW_t^Q \\ &\quad + S_t \left[\left(e^{J_t^Q} - 1\right) dN_t^Q - \mathbb{E}^Q\left(\left(e^{J^Q} - 1\right) \,|\, \mathcal{F}_t\right) \lambda_t^Q dt\right] \end{aligned} \tag{25}$$

4.2 Option pricing

Under the risk neutral measure, the stock price is led by the geometric jump diffusion process presented in equation (25). According to the Itô's lemma, the log return $X_t = \log \frac{S_t}{S_0}$ under Q is then equal to

$$X_t = \int_0^t r - \frac{1}{2}\sigma^2 ds - \mathbb{E}^Q\left(\left(e^{J^Q}-1\right) \mid \mathcal{F}_t\right)\lambda_s^Q ds + \int_0^t \sigma\, dW_t^Q + \sum_{i=1}^{N_t} J_i^Q.$$

The remainder of this section focuses on European options written on the stock log-return. If T is the expiry date, the option payoff is denoted by $V(T, X_T)$. For call and put options with a strike price K, the payoffs are respectively defined by $V(T, X_T) = [S_0 e^{X_T} - K]_+$ and $V(T, X_T) = [K - S_0 e^{X_T}]_+$. The option price is equal to the expected discounted cash-flow. This is also the product of a discount factor and of the integral of the payoff, weighted by the density of $X_T|\mathcal{F}_t$. If this density is denoted by $f_{X_T|\mathcal{F}_t}(x)$, the option value is given by:

$$\mathbb{E}^Q\left(e^{-r(T-t)}V(T, X_T) \mid \mathcal{F}_t\right) = e^{-r(T-t)}\int_{-\infty}^{+\infty} V(T, x)\, f_{X_T|\mathcal{F}_t}(x)dx. \tag{26}$$

The density function does not admit any closed form expression but can be approached numerically by inverting the mgf of the log return, that is provided in proposition 5, by the next discrete Fourier transform (DFT):

Proposition 12: *Let M be the number of steps used in the DFT and let $\Delta_x = \frac{2x_{max}}{M-1}$ be the discretization step. We denote $\Delta_z = \frac{2\pi}{M\,\Delta_x}$ and*

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

for $j = 1 \dots M$. The values of $f_{X_t}(.)$ at points $x_k = -\frac{M}{2}\Delta_x + (k-1)\Delta_x$ are approached by the sum

$$f_{X_T|\mathcal{F}_t}(x_k) \approx \frac{2}{M\,\Delta_x} Re\left(\sum_{j=1}^{M} I_j \varphi\left(i\, z_j, X_t, \lambda_t^Q, \theta_t^Q\right)(-1)^{j-1} e^{-i\frac{2\pi}{M}(j-1)(k-1)}\right), \tag{27}$$

where $I_j = \frac{1}{2}1_{\{j=1\}} + 1_{\{j\neq 1\}}$ and $\varphi\left(i\, z_j, X_t, \lambda_t^Q, \theta_t^Q\right)$ is the mgf of X_T conditionally to $\mathcal{F}_t$.

For a proof, the reader may refer to Hainaut (2016). Once that the density is approached numerically, the option is evaluated by discretizing the integral in equation (26)

$$\mathbb{E}^Q\left(e^{-r(T-t)}V(T, X_T) \mid \mathcal{F}_t\right) \approx e^{-r(T-t)} \sum_{k=1}^{M} V(T, x_k)\, f_{X_T|\mathcal{F}_t}(x_k)\,\Delta_x$$

This method is slightly different from the one proposed by Carr and Madan (1999) who calculate directly the option value by DFT, and for a serie of strike prices.

4.3 Numerical illustration

The first graph of figure 6 shows the surface of implied volatilities for European call options on the S&P 500. The maturities range from 2 weeks to 3 months and strike prices run from 90% to 110% of the spot S&P value. The parameters used for this calculation are these obtained by the econometric calibration. The intensity and its mean reversion level are set to their asymptotic averages, $\lambda_0 = \mathbb{E}(\lambda_\infty)$ and $\theta_0 = \mathbb{E}(\theta_\infty)$. For short term maturities, we observe a pronounced smile of volatilities, slightly asymmetric. For longer maturities, the smile is flatter and around 25%. The sensitivity of the volatility surface to parameters is studied in the next three subplots of figure 6. Increasing the γ shifts up the volatility surface by a few percents. Reducing the speeds of mean reversion or increasing the feedback parameters η and δ both raises implied volatilities. The smile is clearly more sensitive to modifications of the dynamics of λ_t then to these related to θ_t. The two last graphs of figure 6, illustrate the influence of the initial values λ_0 and θ_0 on the short and medium term smiles. Shifting up λ_0 raises the instantaneous probability of observing a train of new jumps, and then increases implied volatilities for all maturities. On the other hand, increasing of θ_0 mainly affects medium term implied volatilties and has nearly no impact on the 1 month smile. Notice that the asymmetry of the smile is mainly controlled by the parameters defining the double exponential jumps.

Insert Figure 6

5 Conclusion

This paper introduces a new category of multifactor self-excited jump diffusion processes in which the occurrence of a jump increases the frequency of future jumps and the level to which it reverts. This model is reformulated as a single factor self-excited jump process with a kernel function that is a mixture of exponential functions. In this framework, we establish closed and semi-closed form expressions for the mean and variance of the intensity, for the moment generating of log returns and for the expected realized variance. We also introduce a class of changes of measure preserving the dynamics of the process under the risk neutral measure. We investigate empirical issues related to the filtering by sequential Monte Carlo of state variables and the econometric calibration by a peak over threhold method. Finally, we show that our process is usable for option pricing and yields realistic smile of volatilities.

The model is fitted to the time serie of S&P 500 daily returns and empirical results corroborate the existence of a relation between the mean reversion level of the jump intensity, and jumps. In particular, the kernel function displays a higher resilience to shocks at long term than the one of the single factor model. This analysis also reveals that self-excitation is induced by positive and negative jumps. From an economic point of view, this means that when the stock market is shacked by a movement of large amplitude, investors should both anticipate an increase of the risk of immediate jump occurrences but also adjust their long term expectation about the frequency of jumps.

6 Appendix

6.1 Exponential and double exponential jumps

The double exponential distribution has certainly been popularized in Finance by the work of Kou (2002). Hereafter we recall a number of well known results. Notice that our notations are slightly different from those used by Kou (2002). A double exponential distributed random variable J may take positive or negative values. Its probability density function (defined on $\mathbb{R}$) is given by

$$\nu(z) = p\rho^+ e^{-\rho^+ z} 1_{\{z \geq 0\}} - (1-p)\,\rho^- e^{-\rho^- z} 1_{\{z<0\}}$$

while the associated cumulative distribution function is

$$\mathbb{P}[J \leq z] = (1-p)\,e^{-\rho^- z} 1_{\{z \leq 0\}} + \left[(1-p) + p\left(1 - e^{-\rho^+ z}\right)\right] 1_{\{z>0\}}.$$

Consequently, this distribution depends on three parameters $\rho^+ \in \mathbb{R}^+$, $\rho^- \in \mathbb{R}^-$ and $p \in (0,1)$ where p (*resp.* $(1-p)$) stands for the probabilities of observing an upward (*resp.* downward) exponential jump and $\frac{1}{\rho^+}$ (*resp.* $\frac{1}{\rho^-}$) gives the average positive (*resp.* negative) size jump. When only unidirectional jumps are considered, all developments remains valid with respectively $p = 1$ or $p = 0$ for positive and negative exponential jumps. The expected value of the size of jump (J) is the weighted sum of these average sizes i.e. $\mathbb{E}(J) = p\frac{1}{\rho^+} + (1-p)\frac{1}{\rho^-}$. The expected value of the absolute value of the size of jump ($|J|$) is $E(|J|) = p\frac{1}{\rho^+} + (1-p)\frac{1}{|\rho^-|} \equiv \mu_J$. An important expression for our research is the second moment of these two random variables

$$E\left(J^2\right) = E\left(|J|^2\right) = p\frac{2}{(\rho^+)^2} + (1-p)\frac{2}{(\rho^-)^2}.$$

Finally, one has

$$\psi(z_1, z_2) := \mathbb{E}\left(e^{z_1 J + z_2 |J|}\right) = p\frac{\rho^+}{\rho^+ - (z_1 + z_2)} + (1-p)\,\frac{\rho^-}{\rho^- - (z_1 - z_2)} \tag{28}$$

if $(z_1 + z_2) < \rho^+$ and $(z_1 - z_2) > \rho^-$ (see Hainaut (2016)) so that $\mathbb{E}\left(e^J\right) = \psi(1,0) = p\frac{\rho^+}{\rho^+ - 1} + (1-p)\,\frac{\rho^-}{\rho^- - 1}$. On the other hand, we have the following useful expressions for the calculation of the expected realized variance and volatility.

$$\mathbb{E}\left(\left(e^J - 1\right)^2\right) = p\left(\frac{\rho^+}{\rho^+ - 2} - \frac{\rho^+ + 1}{\rho^+ - 1}\right) + (1-p)\left(\frac{\rho^-}{\rho^- - 2} - \frac{\rho^- + 1}{\rho^- - 1}\right) \tag{29}$$

and

$$\begin{aligned}\mathbb{E}\left(\left(e^J - 1\right)^4\right) &= p\left(\frac{\rho^+}{\rho^+ - 4} - 4\frac{\rho^+}{\rho^+ - 3} + 6\frac{\rho^+}{\rho^+ - 2} - \frac{3\rho^+ + 1}{\rho^+ - 1}\right) \\ &\quad + (1-p)\left(\frac{\rho^-}{\rho^- - 4} - 4\frac{\rho^-}{\rho^- - 3} + 6\frac{\rho^-}{\rho^- - 2} - \frac{3\rho^- + 1}{\rho^- - 1}\right).\end{aligned}$$

6.2 Binomial jumps

When J has a binomial distribution, its probability density function (defined on $\mathbb{R}$) is given by

$$\nu(z) = p\frac{1}{\rho^+}\delta_{\{z=\rho^+\}} + (1-p)\frac{1}{\rho^-}\delta_{\{z=\rho^-\}}$$

the distribution depends on three parameters $\rho^+ \in \mathbb{R}^+$, $\rho^- \in \mathbb{R}^-$ and $p \in (0,1)$ where p (*resp.* $(1-p)$) stands for the probabilities of observing an upward (*resp.* downward) constant jump of size $\frac{1}{\rho^+}$ (*resp.* $\frac{1}{\rho^-}$) . The expected value of the size of jump (J) is the weighted sum of these average sizes i.e. $\mathbb{E}(J) = p\frac{1}{\rho^+} + (1-p)\frac{1}{\rho^-}$. The expected value of the absolute value of the size of jump ($|J|$) is $E(|J|) = p\frac{1}{\rho^+} + (1-p)\frac{1}{|\rho^-|} \equiv \mu_J$ and

$$E\left(J^2\right) = E\left(|J|^2\right) = p\frac{1}{(\rho^+)^2} + (1-p)\frac{1}{(\rho^-)^2}.$$

The mgf of jumps and of their absolute value is given by

$$\psi(z_1, z_2) := \mathbb{E}\left(e^{z_1 J + z_2 |J|}\right) = pe^{(z_1+z_2)\frac{1}{\rho^+}} + (1-p)\,e^{(z_1-z_2)\frac{1}{\rho^-}} \tag{30}$$

whereas we have the follwing useful expressions for the calculation of the expected realized variance and volatility:

$$\mathbb{E}\left(\left(e^J-1\right)^2\right) = p\left(\left(e^{\frac{1}{\rho^+}}-1\right)^2\right) + (1-p)\left(\left(e^{\frac{1}{\rho^-}}-1\right)^2\right),$$

$$\mathbb{E}\left(\left(e^J-1\right)^4\right) = p\left(\left(e^{\frac{1}{\rho^+}}-1\right)^4\right) + (1-p)\left(\left(e^{\frac{1}{\rho^-}}-1\right)^4\right).$$

6.3 Proofs

Proof of proposition 1: Integrating equation (3) leads to the following expression for λ_t

$$\lambda_t = \int_0^t \alpha e^{-\alpha(t-s)}\theta_s ds + e^{-\alpha t}\lambda_0 + \eta \int_0^t e^{-\alpha(t-s)} dL_s.$$

Now insert the expression (5) of θ_s in this relation. One obtains

$$\lambda_t = \int_0^t \alpha e^{-\alpha(t-s)} \left(\gamma + e^{-\beta s}(\theta_0 - \gamma) + \delta \int_0^s e^{-\beta(s-u)} dL_u\right) ds + e^{-\alpha t}\lambda_0 + \eta \int_0^t e^{-\alpha(t-s)} dL_s$$

or

$$\begin{aligned}\lambda_t &= \gamma \int_0^t \alpha e^{-\alpha(t-s)} ds + (\theta_0 - \gamma) \int_0^t \alpha e^{-\alpha(t-s)} e^{-\beta s} ds + \int_0^t \alpha\delta e^{-\alpha(t-s)} \int_0^s e^{-\beta(s-u)} dL_u ds \\ &\quad + e^{-\alpha t}\lambda_0 + \eta \int_0^t e^{-\alpha(t-s)} dL_s\,.\end{aligned}$$

Then, assuming $\alpha \neq \beta$, changing the order of integration and integrating give

$$\begin{aligned}\lambda_t &= \gamma\left(1 - e^{-\alpha t}\right) + (\theta_0 - \gamma)\frac{\alpha}{(\alpha - \beta)}\left(e^{-\beta t} - e^{-\alpha t}\right) \\ &\quad + e^{-\alpha t}\lambda_0 + \int_0^t \frac{\alpha\delta}{\alpha - \beta}\left(e^{\beta(u-t)} - e^{\alpha(u-t)}\right) + \eta e^{-\alpha(t-u)}\, dL_u\end{aligned}$$

as expected.

If now $\alpha = \beta$, then

$$\begin{aligned}\lambda_t &= \int_0^t \alpha e^{-\alpha(t-s)}\left(\gamma + e^{-\alpha s}(\theta_0 - \gamma) + \delta \int_0^s e^{-\alpha(s-u)} dL_u\right) ds + e^{-\alpha t}\lambda_0 + \eta \int_0^t e^{-\alpha(t-s)} dL_s, \\ &= \gamma + (\lambda_0 - \gamma) e^{-\alpha t} + \eta \int_0^t e^{-\alpha(t-s)} dL_s + \alpha t (\theta_0 - \gamma) e^{-\alpha t} + \delta\alpha \int_0^t e^{-\alpha(t-u)} (t-u)\, dL_u.\end{aligned}$$

□

Proof of proposition 2: As $\mathbb{E}(dL_s|\mathcal{F}_0) = \mathbb{E}(|J|\,|\mathcal{F}_0) \times \mathbb{E}(\lambda_{s-}|\mathcal{F}_0)$, we have from the equation (5) that

$$\mathbb{E}(\theta_t|\mathcal{F}_0) = \gamma + e^{-\beta t}(\theta_0 - \gamma) + \delta\mu_J \int_0^t e^{-\beta(t-s)} \mathbb{E}(\lambda_{s-}|\mathcal{F}_0)\, ds.$$

If we derive this last expression with respect to time, we find that $\mathbb{E}(\theta_t|\mathcal{F}_0)$ is solution of an ODE:

$$\begin{aligned}\frac{\partial}{\partial t}\mathbb{E}(\theta_t|\mathcal{F}_0) &= -\beta e^{-\beta t}(\theta_0 - \gamma) + \delta\mu_J \mathbb{E}(\lambda_t|\mathcal{F}_0) - \beta\delta\mu_J \int_0^t e^{-\beta(t-s)} \mathbb{E}(\lambda_{s-}|\mathcal{F}_0)\, ds \\ &= \delta\mu_J \mathbb{E}(\lambda_t|\mathcal{F}_0) - \beta\, \mathbb{E}(\theta_t|\mathcal{F}_0) + \gamma\beta\,. \qquad (31)\end{aligned}$$

On another hand, the expectation of λ_t is also given by the relation

$$\mathbb{E}\left(\lambda_t|\mathcal{F}_0\right) = \alpha \int_0^t e^{-\alpha(t-s)}\mathbb{E}\left(\theta_s|\mathcal{F}_0\right) ds + e^{-\alpha t}\lambda_0 + \eta\mu_J \int_0^t e^{-\alpha(t-s)}\mathbb{E}\left(\lambda_{s-}|\mathcal{F}_0\right) ds.$$

And if we derive this last expression with respect to time, we obtain the following ODE for $\mathbb{E}\left(\lambda_t|\mathcal{F}_0\right)$:

$$\begin{aligned}
\frac{\partial}{\partial t}\mathbb{E}\left(\lambda_t|\mathcal{F}_0\right) &= \alpha\mathbb{E}\left(\theta_t|\mathcal{F}_0\right) - \alpha\int_0^t \alpha e^{-\alpha(t-s)}\mathbb{E}\left(\theta_s|\mathcal{F}_0\right) ds - \alpha e^{-\alpha t}\lambda_0 + \\
&\quad \eta\mu_J\mathbb{E}\left(\lambda_t|\mathcal{F}_0\right) - \alpha\eta\mu_J\int_0^t e^{-\alpha(t-s)}\mathbb{E}\left(\lambda_{s-}|\mathcal{F}_0\right) ds \\
&= \alpha\mathbb{E}\left(\theta_t|\mathcal{F}_0\right) + \left(\eta\mu_J - \alpha\right)\mathbb{E}\left(\lambda_t|\mathcal{F}_0\right) .
\end{aligned}$$

In matrix form, $\mathbb{E}\left(\theta_t|\mathcal{F}_0\right)$ and $\mathbb{E}\left(\lambda_t|\mathcal{F}_0\right)$ are solutions of a system of ODE:

$$\begin{pmatrix} \frac{\partial}{\partial t}\mathbb{E}\left(\theta_t|\mathcal{F}_0\right) \\ \frac{\partial}{\partial t}\mathbb{E}\left(\lambda_t|\mathcal{F}_0\right) \end{pmatrix} = \underbrace{\begin{pmatrix} -\beta & \delta\mu_J \\ \alpha & (\eta\mu_J - \alpha) \end{pmatrix}}_{M} \begin{pmatrix} \mathbb{E}\left(\theta_t|\mathcal{F}_0\right) \\ \mathbb{E}\left(\lambda_t|\mathcal{F}_0\right) \end{pmatrix} + \begin{pmatrix} \gamma\beta \\ 0 \end{pmatrix} \tag{32}$$

Finding a solution requires to determine eigenvalues γ and eigenvectors (v_1, v_2) of the matrix M, multiplying the expectations. These eigenvalues cancel the following determinant of

$$det\begin{pmatrix} -\beta-\gamma & \delta\mu_J \\ \alpha & (\eta\mu_J - \alpha) - \gamma \end{pmatrix} = 0$$

and are roots of a second order polynomial

$$\begin{aligned}
0 &= \left(-\beta-\gamma\right)\left(\left(\eta\mu_J - \alpha\right) - \gamma\right) - \alpha\delta\mu_J \\
&= -\alpha\delta\mu_J - \beta(\eta\mu_J - \alpha) + \gamma\left[\beta - (\eta\mu_J - \alpha)\right] + \gamma^2
\end{aligned}$$

$$0 = -\alpha\delta\mu_J - \beta(\eta\mu_J - \alpha) + \left(\beta - (\eta\mu_J - \alpha)\right)\gamma + \gamma^2 .$$

If the discriminant, Δ, is positive

$$\Delta = \left(\beta - (\eta\mu_J - \alpha)\right)^2 + 4\left(\delta\mu_J\alpha + \beta(\eta\mu_J - \alpha)\right) > 0$$

then the roots γ_1 and γ_2 are given by:

$$\begin{aligned}
\gamma_1 &= -\frac{1}{2}\left(\beta - (\eta\mu_J - \alpha)\right) + \frac{1}{2}\sqrt{\Delta}\,, \\
\gamma_2 &= -\frac{1}{2}\left(\beta - (\eta\mu_J - \alpha)\right) - \frac{1}{2}\sqrt{\Delta}\,.
\end{aligned}$$

The product of the roots is the constant of the second order polynomial

$$\gamma_1\gamma_2 = -\alpha\delta\mu_J - \beta(\eta\mu_J - \alpha) \tag{33}$$

and, by passing, one has

$$
\begin{aligned}
& (\beta+\gamma_2)(\beta+\gamma_1) \\
= & \frac{1}{4}\left([2\beta+((\eta\mu_J-\alpha)-\beta)]^2-\left[(\beta-(\eta\mu_J-\alpha))^2+4(\delta\mu_J\alpha+\beta(\eta\mu_J-\alpha))\right]\right) \\
= & \frac{1}{4}\left(\underline{\underline{4\beta^2}}+4\beta\left(\underline{(\eta\mu_J-\alpha)}-\underline{\underline{\beta}}\right)+\underbrace{((\eta\mu_J-\alpha)-\beta)^2-(\beta-(\eta\mu_J-\alpha))^2}_{=0}\right) \\
& -\delta\mu_J\alpha-\underline{\beta(\eta\mu_J-\alpha)} \\
= & -\delta\mu_J\alpha. \qquad (34)
\end{aligned}
$$

Eigenvectors of the matrix M are orthogonal and such that

$$\begin{pmatrix} -\beta & \delta\mu_J \\ \alpha & (\eta\mu_J-\alpha) \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}=0.$$

Then we infer that

$$\begin{pmatrix} v_1^i \\ v_2^i \end{pmatrix}=\begin{pmatrix} -\delta\mu_J \\ -\beta-\gamma_i \end{pmatrix} : i=1,2$$

Finally, if $D=diag(\gamma_1,\gamma_2)$ and V is the matrix of eigenvectors, the matrix M in equation (32) admits the following decomposition

$$\begin{pmatrix} -\beta & \delta\mu_J \\ \alpha & (\eta\mu_J-\alpha) \end{pmatrix}=VDV^{-1}.$$

If we define

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}=V^{-1}\begin{pmatrix} \mathbb{E}(\theta_t|\mathcal{F}_0) \\ \mathbb{E}(\lambda_t|\mathcal{F}_0) \end{pmatrix}$$

the system (32) can be rewritten as two independent ODE's

$$\begin{pmatrix} \frac{\partial}{\partial t}u_1 \\ \frac{\partial}{\partial t}u_2 \end{pmatrix}=\begin{pmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}+V^{-1}\begin{pmatrix} \gamma\beta \\ 0 \end{pmatrix}.$$

If we introduce the following notation,

$$V^{-1}\begin{pmatrix} \gamma\beta \\ 0 \end{pmatrix}=\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix}$$

the solutions of ODE's are given by

$$
\begin{aligned}
u_1(t) &= \frac{\epsilon_1}{\gamma_1}\left(e^{\gamma_1 t}-1\right)+d_1e^{\gamma_1 t} \\
u_2(t) &= \frac{\epsilon_2}{\gamma_2}\left(e^{\gamma_2 t}-1\right)+d_2e^{\gamma_2 t}
\end{aligned}
$$

where $d=(d_1,d_2)'$ is such that $d=V^{-1}\begin{pmatrix}\theta_0\\ \lambda_0\end{pmatrix}$. Then

$$\begin{pmatrix}u_1\\ u_2\end{pmatrix}=\begin{pmatrix}\frac{1}{\gamma_1}\left(e^{\gamma_1 t}-1\right) & 0\\ 0 & \frac{1}{\gamma_2}\left(e^{\gamma_2 t}-1\right)\end{pmatrix}V^{-1}\begin{pmatrix}\gamma\beta\\ 0\end{pmatrix}+\begin{pmatrix}e^{\gamma_1 t} & 0\\ 0 & e^{\gamma_2 t}\end{pmatrix}V^{-1}\begin{pmatrix}\theta_0\\ \lambda_0\end{pmatrix}$$

and we can conclude. □

Proof of proposition 3: We define the martingale $M_t := \lambda_t - \mathbb{E}(\lambda_t|\mathcal{F}_0)$, then $[\lambda,\lambda]_t = [M,M]_t$ and $[M,M]_t = M_t^2 - 2\int_0^t M_{s-}dM_s$. Then $\mathbb{V}\left(\lambda_t|\mathcal{F}_0\right)=\mathbb{E}\left([M,M]_t|\mathcal{F}_0\right)=\mathbb{E}\left([\lambda,\lambda]_t|\mathcal{F}_0\right)$. The expected quadratic variation of λ, is given by

$$\begin{aligned}
[\lambda,\lambda]_t &= \Bigg[\int_0^t \frac{\alpha\delta}{\alpha-\beta}\left(e^{\beta(u-t)}-e^{\alpha(u-t)}\right)+\eta e^{-\alpha(t-u)}\,dL_u\,,\\
&\qquad \int_0^t \frac{\alpha\delta}{\alpha-\beta}\left(e^{\beta(u-t)}-e^{\alpha(u-t)}\right)+\eta e^{-\alpha(t-u)}\,dL_u\Bigg]_t\\
&= \int_0^t\left(\frac{\alpha\delta}{\alpha-\beta}\left(e^{\beta(u-t)}-e^{\alpha(u-t)}\right)+\eta e^{-\alpha(t-u)}\right)^2|J_u|^2dN_u
\end{aligned} \tag{35}$$

From this last equation and as $\mathbb{E}\left(|J|^2\right)=p\frac{2}{(\rho^+)^2}+(1-p)\frac{2}{(\rho^-)^2}$, we infer that:

$$\mathbb{V}\left(\lambda_t|\mathcal{F}_0\right)=\mathbb{E}(|J|^2)\int_0^t\left(\frac{\alpha\delta}{\alpha-\beta}e^{\beta(u-t)}+\left(\eta-\frac{\alpha\delta}{\alpha-\beta}\right)e^{\alpha(u-t)}\right)^2\mathbb{E}\left(\lambda_u|\mathcal{F}_0\right)du$$

If we inject the expression (12) for $\mathbb{E}\left(\lambda_u|\mathcal{F}_0\right)$ into this last equation, we conclude. □.

Proof of Proposition 4: To lighten notations, we temporarily denote $f=\mathbb{E}\left(e^{\omega_0\Lambda_s+\omega_1\lambda_s+\omega_2\theta_s}\,|\,\mathcal{F}_t\right)$, with $t\leq s$. It is solution of the following Itô's equation:

$$\begin{aligned}
0 &= f_t + f_\Lambda\lambda_t + \alpha(\theta_t-\lambda_t)\,f_\lambda + \beta\left(\gamma-\theta_t\right)f_\theta\\
&\quad +\lambda_t\int_0^{+\infty} f\left(t,\,\lambda_t+\eta\,|z|,\theta_t+\delta\,|z|\right)-f(.)\,d\nu(z)\,.
\end{aligned} \tag{36}$$

In the remainder of this proof, f is assumed to be an exponential affine function of λ_t, θ_t and Λ_t :

$$f = exp\left(A(t,s)+B(t,s)\lambda_t+C(t,s)\theta_t+D(t,s)\Lambda_t\right),$$

where $A(t,s)$, $B(t,s)$, $C(t,s)$ and $D(t,s)$ are functions of time with the terminal conditions $A(s,s)=0$, $B(s,s)=\omega_1$, $C(s,s)=\omega_2$ and $D(s,s)=\omega_0$. Under this assumption, the partial derivatives of f are given by:

$$f_t=\left(\frac{\partial}{\partial t}A(t,s)+\frac{\partial}{\partial t}B(t,s)\lambda_t+\frac{\partial}{\partial t}C(t,s)\theta_t+\frac{\partial}{\partial t}D(t,s)\Lambda_t\right)f,$$

$$f_\Lambda = D(t,s)f\ \ f_\lambda = B(t,s)f : f_\theta = C(t,s)f$$

And the integrand in equation (36) is becomes:

$$\int_0^{+\infty} f\left(t,\ \lambda_t+\eta\,|z|,\theta_t+\delta\,|z|\right)-f(.)\,d\nu(z)=f\left[\psi\left(0,\ B(t,s)\eta+C(t,s)\delta\right)-1\right].$$

Injecting these expressions into the equation (36), leads to the following relation:

$$\begin{aligned}0 &= \left(\frac{\partial}{\partial t}A(t,s)+\beta\gamma C(t,s)\right)+\frac{\partial}{\partial t}D(t,s)\Lambda_t+\theta_t\left(\frac{\partial}{\partial t}C(t,s)+\alpha B(t,s)-\beta C(t,s)\right)\\ &\quad +\lambda_t\left(\frac{\partial}{\partial t}B(t,s)+D(t,s)-\alpha B(t,s)+\left[\psi\left(0,B(t,s)\eta+C(t,s)\delta\right)-1\right]\right)\end{aligned}$$

This relation implies that $D(t,s)=\omega_0$ and that $A(t,s)$, $B(t,s)$, $C(t,s)$ satisfy the system of ODEs:

$$\begin{cases}\frac{\partial}{\partial t}A &= -\beta\gamma C\\ \frac{\partial}{\partial t}B &= \alpha B-\left[\psi\left(0,B\eta+C\delta\right)+\omega_0-1\right]\\ \frac{\partial}{\partial t}C &= -\alpha B+\beta C\end{cases}. \tag{37}$$

□

Proof of Proposition 5: As usual, let us denote $f=\mathbb{E}\left(e^{\omega_1 X_s}\,|\,\mathcal{F}_t\right)$, with $t\leq s$. It is solution of the following Itô's equation for semi martingale:

$$\begin{aligned}0 &= f_t+f_X\left(\mu-\frac{\sigma^2}{2}-\lambda_t\mathbb{E}\left(e^J-1\right)\right)+f_{XX}\frac{\sigma^2}{2}+\alpha(\theta_t-\lambda_t)\,f_\lambda+\beta\left(\gamma-\theta_t\right)f_\theta \\ &\quad +\lambda_t\int_0^{+\infty} f\left(t,X_t+z,\ \lambda_t+\eta\,|z|,\theta_t+\delta\,|z|\right)-f(.)\,d\nu(z)\,.\end{aligned} \tag{38}$$

If f is assumed to be an exponential affine function of λ_t, θ_t, and X_t :

$$f=exp\left(A(t,s)+B(t,s)\lambda_t+C(t,s)\theta_t+D(t,s)X_t\right),$$

where $A(t,s)$, $B(t,s)$, $C(t,s)$, $D(t,s)$ are time dependent functions. The result is proven in a similar way to proposition 4. □

Proof of Proposition 8: Let us denote by Y_t the exponent of M_t:

$$\begin{aligned}Y_t &= \left(\kappa_1(\xi),\ \kappa_2(\xi)\right)\begin{pmatrix}\lambda_t\\ \theta_t\end{pmatrix}+\xi\,L_t-\kappa_3(\xi)t \\ &\quad -\frac{1}{2}\int_0^t\varphi(s)^2ds-\int_0^t\varphi(s)dW_s\end{aligned} \tag{39}$$

Its infinitesimal dynamics is given by

$$\begin{aligned}dY_t &= \kappa_1\alpha\left(\theta_t-\lambda_t\right)dt+\kappa_2\beta\left(\gamma-\theta_t\right)dt\\ &\quad +\left(\kappa_1\eta+\kappa_2\delta+\xi\right)|J|dN_t-\kappa_3 dt-\frac{1}{2}\varphi(t)^2dt-\varphi(t)dW_t\end{aligned}$$

In the remainder of this proof, the random measure of J is noted $\Xi(.)$ and is such that $J = \int_0^\infty \Xi(dz)$. Applying the Ito's lemma for semi-martingales to M_t leads to the next relation:

$$dM_t = M_t dY_t + \frac{1}{2} M_t d\left[Y_t, Y_t\right]_t^c$$
$$+M_t \int_0^\infty \left(e^{(\kappa_1\eta+\kappa_2\delta+\xi)|z|} - 1 - (\kappa_1\eta + \kappa_2\delta + \xi)\,|z| \right) \Xi(dz) dN_t$$

that is developed as follows:

$$\begin{aligned} dM_t &= M_t \left((\kappa_1\alpha - \kappa_2\beta)\,\theta_t + \kappa_2\beta\gamma - \kappa_3 \right) dt - \frac{1}{2}\varphi(t)^2 M_t dt + \frac{1}{2} M_t \varphi(t)^2 dt \\ & -\varphi(t) M_t dW_t - M_t \lambda_t \left(\kappa_1\alpha - \int_0^\infty \left(e^{(\kappa_1\eta+\kappa_2\delta+\xi)|z|} - 1 \right) \nu(dz) \right) dt \\ & +M_t \int_0^\infty \left(e^{(\kappa_1\eta+\kappa_2\delta+\xi)|z|} - 1 \right) \left(\Xi(dz) dN_t - \lambda_t \nu(dz) dt \right) \end{aligned}$$

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

$$\begin{cases} \kappa_1\alpha - \kappa_2\beta & = 0 \\ \kappa_2\beta\gamma - \kappa_3 & = 0 \\ \kappa_1\alpha - \int_0^\infty \left(e^{(\kappa_1\eta+\kappa_2\delta+\xi)|z|} - 1 \right) \nu(dz) & = 0 \end{cases}$$

□

Proof of proposition 9: If Y_t is the exponent of M_t, as defined by equation (39) and if we note $\psi^b = \psi(0, \kappa_1\eta + \kappa_2\delta + \xi)$, the expectation under Q of the mgf of λ_T^Q and θ_T^Q is equal to

$$\begin{aligned} \mathbb{E}^Q \left(e^{\omega_1 \lambda_T^Q + \omega_2 \theta_T^Q} | \mathcal{F}_t \right) &= \mathbb{E}\left(e^{Y_T - Y_t + \omega_1 \psi^b \lambda_T + \omega_2 \psi^b \theta_T} | \mathcal{F}_t \right) \\ &= e^{-Y_t} \mathbb{E}\left(e^{Y_T + \omega_1 \psi^b \lambda_T + \omega_2 \psi^b \theta_T} | \mathcal{F}_t \right) . \end{aligned}$$

If $f(.)$ denotes $\mathbb{E}\left(e^{Y_T + \omega_1 \psi^b \lambda_T + \omega_2 \psi^b \theta_T} | \mathcal{F}_t \right)$, according to the Itô's lemma, it solves the differential equation:

$$\begin{aligned} 0 = f_t &+ \left(\kappa_1\alpha\,(\theta_t - \lambda_t) + \kappa_2\beta\,(\gamma - \theta_t) - \kappa_3 \right) f_Y + \\ & \alpha\,(\theta_t - \lambda_t)\, f_\lambda + \beta\,(\gamma - \theta_t)\, f_\theta - \frac{1}{2}\varphi(t)^2\, f_Y\, dt + \frac{1}{2}\varphi(t)^2 f_{YY} + \\ & \lambda_t \int_{-\infty}^{+\infty} f\left(t, \lambda_t + \eta|z|\,, \theta_t + \delta|z|\,, Y_t + (\kappa_1\eta + \kappa_2\delta + \xi)\,|z| \right) - f\, d\nu(z)\,. \end{aligned} \tag{40}$$

As usual, we assume that $f(.)$ is an exponential affine function of state variables:

$$f = \exp\left(A(t,T) + \psi(0, \kappa_1\eta + \kappa_2\delta + \xi)\left(B(t,T)\lambda_t + C(t,T)\theta_t\right) + D(t,T)Y_t\right) ,$$

and then

$$f_t = \left(\frac{\partial}{\partial t}A + \psi^b \lambda_t \frac{\partial}{\partial t}B + \psi^b \frac{\partial}{\partial t}C\,\theta_t + \frac{\partial}{\partial t}D\,Y_t\right) f$$

$$f_Y = D\,f \quad f_{YY} = D^2\,f \quad f_\lambda = B\psi^b\,f \quad f_\theta = C\psi^b\,f$$

Inserting these expressions into the equation (40), leads to the following relation (after grouping terms):

$$\begin{aligned}
0 &= \frac{\partial}{\partial t}A + \kappa_2\beta\gamma D - \kappa_3 D + \beta\gamma\psi^b C - \frac{1}{2}\varphi(t)^2 D f + \frac{1}{2}\varphi(t)^2 D^2 f \\
&+\lambda_t\left(\psi^b\frac{\partial}{\partial t}B - \kappa_1\alpha D - \alpha\psi^b B + \int_0^{+\infty}\left[e^{B\psi^b\eta|z| + C\psi^b\delta|z| + D(\kappa_1\eta+\kappa_2\delta+\xi)|z|} - 1\right] d\nu(z)\right) \\
&+\theta_t\left(\psi^b\frac{\partial}{\partial t}C + \kappa_1\alpha D - \kappa_2\beta D + \alpha\psi^b B - \beta\psi^b C\right) + Y_t\left(\frac{\partial}{\partial t}D\right)
\end{aligned}$$

we infer that $D(t,s) = 1$ as $\frac{\partial}{\partial t}D(t,s) = 0$. And we get that

$$\begin{aligned}
0 &= \frac{\partial}{\partial t}A + \kappa_2\beta\gamma - \kappa_3 + \beta\gamma\psi^b C \\
0 &= \psi^b\frac{\partial}{\partial t}B - \kappa_1\alpha - \alpha\psi^b B + \int_0^{+\infty}\left[e^{B\psi^b\eta|z| + C\psi^b\delta|z| + (\kappa_1\eta+\kappa_2\delta+\xi)|z|} - 1\right] d\nu(z) \\
0 &= \psi^b\frac{\partial}{\partial t}C + \kappa_1\alpha - \kappa_2\beta + \alpha\psi^b B - \beta\psi^b C : .
\end{aligned}$$

Using conditions (21), this system is simplified as follows:

$$\begin{aligned}
0 &= \frac{\partial}{\partial t}A + \beta\gamma\psi^b C \\
0 &= \frac{\partial}{\partial t}B - \alpha B + \left[\frac{\psi\left(0, B\psi^b\eta + C\psi^b\delta + (\kappa_1\eta + \kappa_2\delta + \xi)\right)}{\psi^b} - 1\right] \\
0 &= \frac{\partial}{\partial t}C + \alpha B - \beta C
\end{aligned}$$

and we can conclude by comparison with the results of proposition 2. □

Proof of proposition 10: By construction, the moment-generating function for jumps under the risk-neutral measure is the ratio

$$\psi^Q(z, 0) = \frac{\psi\left(z, (\kappa_1\eta + \kappa_2\delta + \xi)\right)}{\psi\left(0, \kappa_1\eta + \kappa_2\delta + \xi\right)}.$$

If we denote $\alpha = (\kappa_1 \eta + \kappa_2 \delta + \xi)$, the numerator and denominator in this equation are given by

$$\begin{aligned}
\psi(z, \alpha) &= \frac{p\rho^+ (\rho^- + \alpha - z) + (1-p)\rho^- (\rho^+ - \alpha - z)}{(\rho^+ - \alpha - z)(\rho^- + \alpha - z)} \\
\psi(0, \alpha) &= \frac{p\rho^+ (\rho^- + \alpha) + (1-p)\rho^- (\rho^+ - \alpha)}{(\rho^+ - \alpha)(\rho^- + \alpha)}.
\end{aligned}$$

Then, since

$$\psi^Q(z, 0) = \frac{\frac{p\rho^+\rho^{-Q}}{(p\rho^+\rho^{-Q} + (1-p)\rho^-\rho^{+Q})}\left(\rho^{-Q} - z\right)\rho^{+Q} + \frac{(1-p)\rho^-\rho^{+Q}}{(p\rho^+\rho^{-Q} + (1-p)\rho^-\rho^{+Q})}\rho^{-Q}\left(\rho^{+Q} - z\right)}{(\rho^{+Q} - z)(\rho^{-Q} - z)},$$

the proof is complete. □

6.4 Econometric methodology

Among the most popular approach for calibration in the financial literature, we find the method of simulated moments (Duffie and Singleton (1993)), indirect inference methods (Gourieroux, Monfort, and Renault (1993)) and the efficient method of moments (EMM) (Gallant and Tauchen (1996)). We opt instead for an asymmetric "peaks over threshold" (POT) procedure similar to the one of Embrechts et al. (2011). This method is robust and less time consuming than methods based on simulations or Bayesian Learning as used by Fulop et al. (2015). Even if the biais of estimation is more important with the POT procedure, this is sufficient for our purpose which is to demonstrate the validity of the bifactor model.

The discrete record of T observations of log returns, equally spaced with a lag Δ of one day of trading is noted $\{x_{1,}x_1, x_2, ..., x_T\}$. Whether this return is above or below some thresholds, it is likely that a jump occurred. These thresholds, denoted $g(\alpha_1)$ and $g(\alpha_2)$ depend on the lag between observations and of confidence levels, α_1 α_2. To define them, we fit by log likelihood maximization, a pure Gaussian process : $x_i \sim \mu\Delta + \sigma W_\Delta$. And if $\Phi(.)$ denotes the pdf of a standard normal, $g(\alpha_1)$, $g(\alpha_2)$ are set to the α_1 and α_2 percentiles of the Brownian term: $g(\alpha_i) = \sigma\sqrt{\Delta}\Phi^{-1}(\alpha_i)$. When a jump is detected, the dynamics of the stocks index is approached by:

$$\begin{cases}(x_i - \mu\Delta) \sim J_i & (x_i - \mu\Delta) > g(\alpha_1)\, or\, (x_i - \mu\Delta) < g(\alpha_2)\end{cases}$$

Levels of confidence, α_1 and α_2 are optimized such that the skewness and the kurtosis of x_i for periods without jump are close to these of a normal distribution. For the S&P, we find that α_1 and α_2 are respectively equal to 94% and 91%. The skewness and kurtosis of returns for days without detected jumps are equal to 0.047 and 3.28. The volatilities of the sample from which we withdraw positive, negative and both type of jumps are 18%, 16% and 12%. Once that jumps are detected, the sample paths of θ_t and λ_t for a given set of parameters are approached by:

$$\begin{aligned}\Delta\theta_i &= \beta(\gamma - \theta_{i-1})\Delta + \delta\, J_i\, I_{jump\, at\, t_i}\\ \Delta\lambda_i &= \alpha(\theta_i - \lambda_{i-1})\Delta + \eta\; J_i\, I_{jump\, at\, t_i}\end{aligned}$$

Then, jumps and intensities are calibrated next by maximizing two log likelihoods: one for the distribution of jumps and one for the dynamics of λ_t. E.g. for the DSEJD,

$$\begin{cases}(\rho^-, \rho^+, p) & = \arg\max \sum_{i=1}^{n} \log \nu\,(x_i \mid \rho^-, \rho^+, p)\, I_{jump\, at\, t_i}\\ (\alpha, \eta, \beta, \delta, \gamma, \lambda_0) & = \arg\max \sum_{i=1}^{n} \log P\,(\Delta_i N \mid \lambda_i h)\, I_{jump\, at\, t_i}\end{cases}$$

The log-likelihood of the whole model is next estimated with the particle filter, introduced in the next section. An inherent problem of particle filters is that the estimate of the likelihood is not continuous as a function of parameters. From a practical viewpoint, maximising the resulting rough surface to calibrate the process is then too computationally expensive and uncertain. This is motivate the choice of the peak over threshold method, which is a more robust alternative.

6.4.1 Filtering state variables

Our estimation approach relies on a time-discretization of our SEDJ model characterized by equations (1), (3) and (5). Denote by Δ the length of the time interval. The *ex ante* continuously compounded return (over the period Δ) at time $t = j\Delta$ defined by $X_j = \ln S_{(j+1)\Delta} - \ln S_{j\Delta}$ then satisfies the following equation in discrete time

$$X_j = \underbrace{\left(\mu - \frac{\sigma^2}{2} - \lambda_j \mathbb{E}\left(e^J - 1\right)\right)}_{\mu_j} \Delta + \sigma\sqrt{\Delta}\varepsilon_j + \mathbf{\Delta} L_j^{'} \tag{41}$$

where ε_j stands for a standard normal random variable and $\mathbf{\Delta} L_j^{'} = \sum_{k=N_{j\Delta}}^{N_{(j+1)\Delta}} J_k$. Here $N_{(j+1)\Delta} - N_{j\Delta}$ is distributed as a Poisson random variable with parameter $\lambda_{j-1}\Delta$. The Euler approximation of equations (3) and (5) then provide the discretized dynamics of the latent variables $\lambda = (\lambda_t)_t$ and $\theta = (\theta_t)_t$ given by

$$\lambda_j = \lambda_{j-1} + \quad \alpha(\theta_{j-1} - \lambda_{j-1})\Delta + \eta\, \mathbf{\Delta} L_j \tag{42}$$

$$\theta_j = \theta_{j-1} + \quad \beta\left(\gamma - \theta_{j-1}\right)\Delta + \delta\, \mathbf{\Delta} L_j \tag{43}$$

where $\mathbf{\Delta} L_j = \sum_{k=N_{j-1}}^{N_j} |J_k|$. It is worth noticing that, in our applications, Δ is one day the duration between two consecutive observations and that, in such a $\Delta-$long period, there is a very small probability that more than one jump occurs (this probability is equal to $\frac{(\lambda_{j-1}\Delta)^2}{2} e^{-\lambda_{j-1}\Delta}$). For very short Δ, one may consider $\mathbf{\Delta} L_j^{'} = J\xi_j$ and $\mathbf{\Delta} L_j = |J|\,\xi_j$, where $\xi_j = 1_{\{N_{(j+1)\Delta} - N_{j\Delta} = 1\}}$ is a Bernoulli random variable with probability $\lambda_{j-1}\Delta$.

Notice that, at this stage, the model parameters are assumed to be known. Denote $v_j = (\lambda_j, \theta_j, N_j, \mathbf{\Delta} L_j, \mathbf{\Delta} L_j^{'})$ the 'particle' that puts together all necessary information about the jump process at time $t = j\Delta$.

The above system of equations admits a useful state-space representation, where the equation (41) provides a measurement equation or system (the 'space') that defines the relationship between the (possibly observed) return and the underlying state variables. The vector $v_j = (\lambda_j, \theta_j, N_j, \mathbf{\Delta} L_j, \mathbf{\Delta} L_j^{'})$ can help finding the transition system (the 'state') that describes the dynamics of the state variables. This dynamics depends intimately on the equations (42) and (43).

Denote the sample of observed continuously compounded returns by $\{x_1, x_2, ..., x_n\}$ and $\mathcal{G}_j$ the set of information associated to the subset $\{x_1, x_2, ..., x_j\}$, then $\mathcal{G} = (\mathcal{G}_j)_j$ forms a filtration.

Conditional to information contained in v_j, the return density $p(x_j|v_j)$ is Gaussian $p(x_j|v_j) = \mathcal{N}\left(\mu_j\Delta - \mathbf{\Delta} L_j^{'}\,,\, \sigma\sqrt{\Delta}\right)$. On another side, it is possible to simulate the transition density $p(v_{j+1}\,|\,v_j)$ with equations (42) and (43). The density of v_0 is $p(v_0)$ and the posterior distribution of v_j is denoted by $p(v_j\,|\,y_{1:j})$. As $P(A|B) = \frac{P(A \cup B)}{P(B)}$, this posterior distribution can be rewritten as follows

$$p(v_j\,|\,x_{1:j}) = \frac{p(x_{1:j}\,,\, v_j\,)}{p(x_{1:j})}, \tag{44}$$

and according to the Bayes' rule, the denominator satisfies the equality:

$$p(x_{1:j}) = p(x_{1:j-1}, x_j) = p(x_j \,|\, x_{1:j-1})p(x_{1:j-1})$$

On another hand, the numerator of equation (44) is developped as follows

$$p(x_{1:j}\,,\,v_j\,) = p(x_j \,|\, v_j)p(v_j \,|\, x_{1:j-1})p(x_{1:j-1}).$$

Finally, we obtain the expression for the posterior distribution is:

$$p(v_j \,|\, x_{1:j}) = \frac{p(x_j \,|\, v_j)}{\int p(x_j|v_j)p(v_j \,|\, x_{1:j-1})dv_j} p(v_j \,|\, x_{1:j-1}) \tag{45}$$

where

$$p(v_j \,|\, x_{1:j-1}) = \int p(v_j \,|\, v_{j-1})p(v_{j-1} \,|\, x_{1:j-1})dv_{j-1} \tag{46}$$

To summarize, the calculation of $p(\lambda_j, \theta_j \,|\, x_{1:j})$ is done in two steps, close to these used in the Kalman Filter. The first one is a prediction step in which we estimate $p(v_j \,|\, x_{1:j-1})$ by the relation (46). In the correction step, we next calculate the probabilities $p(v_j \,|x_{1:j})$ using the equation (45). In practice, the integral in the prediction step is replaced by a Monte Carlo simulation, of N particles, noted $v_j^{(i)} = (\lambda_j^{(i)}, \theta_j^{(i)}, \mathbf{\Delta} L_j^{(i)}, \mathbf{\Delta} L_j^{\prime(i)})$ for $i = 1, \ldots, N$. And the structure of the particle filter algorithm is the following:

1. **Initial step**: draw N values of $v_0^{(i)}$ for $i = 1, \ldots, N$, from an initial distribution $p(v_0)$

2. For $j = 1 \,:\, T$
   Prediction step: draw a sample $\Delta L_j^{(i)}$ and $\Delta L_j^{\prime(i)}$ and update $\lambda_j^{(i)}$, $\theta_j^{(i)}$ using the relations

$$\begin{aligned} \lambda_j^{(i)} &= \lambda_{j-1}^{(i)} + \alpha(\theta_{j-1}^{(i)} - \lambda_{j-1}^{(i)})\Delta + \eta \mathbf{\Delta} L_j^{(i)} \\ \theta_j^{(i)} &= \theta_{j-1}^{(i)} + \beta \left(\gamma - \theta_{j-1}^{(i)}\right) dt + \delta \mathbf{\Delta} L_j^{(i)} \end{aligned}$$

   Correction step: the particle $v_j^{(i)}$ has a probability of occurrence equal to $w_j^{(i)} = \frac{p(x_j \,|\, v_j)}{\sum_{i=1:N} p(x_j \,|\, v_j)}$ where $p(x_j \,|\, v_j) \sim \mathcal{N}\left(\mu_j^{(i)}\Delta - \mathbf{\Delta} L_j^{\prime(i)}\,,\, \sigma\sqrt{\Delta}\right)$.
   Resampling step: resample with replacement N particles according to the importance weights $w_j^{(i)}$. The new importance weights are set to $w_j^{(i)} = \frac{1}{N}$.

Notice that there exists smoothing procedures as the Forward-Backward Smoother (FBS) or Maximum A Posteriori Smoother (MAP). However these procedures are time consuming and as showed by Ncube (2009) in his PhD thesis, the gain of accuracy is sometimes limited.

References

[1] Andersen, T., Benzoni, L., Lund, J., 2002. An empirical investigation of continuous-time equity return models. *Journal of Finance* 57, 1239-1284.

[2] Ait-Sahalia, Y., Cacho-Diaz, J., Laeven, R.J.A. 2015. Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, 117, 585-606.

[3] Ait-Sahalia, R. J. A. Laeven, and L. Pelizzon. 2014. Mutual excitation in eurozone sovereign CDS. *Journal of Econometrics*, 183(2):151–167.

[4] At-Sahalia, Y., Karaman, M., Mancini, L. 2014. The term structure of variance swaps and risk premia. *working paper*, Princeton University.

[5] Bacry E., Delattre S.,Hoffmann M., Muzy J. F. 2013. Modelling microstructure noise with mutually exciting point processes. *Quantitative Finance*, 13(1), 65-67.

[6] Bates, D., 2000. Post-87 crash fears in S&P500 futures options. *Journal of Econometrics* 94, 181-238.

[7] Bauwens, L. and N. Hautsch, 2009: Modelling Financial High Frequency Data Using Point Processes, in: T. Mikosch, J.-P. Kreiß, R. A. Davis, T. G. Andersen (Eds.), *Handbook of Financial Time Series*, Springer Berlin, 953–979.

[8] Bowsher C. 2007. Modelling security market events in continuous time: Intensity based, multivariate point process models, *Journal of Econometrics* 141(2), 876–912

[9] Brockhaus O., Long D. 2000. Volatility swaps made simple. *Risk*, Jan., 92-96.

[10] Broadie, M., Jain, A. 2008. The effect of jumps and discrete sampling on volatility and variance swaps. *International Journal of Theoretical and Applied Finance*, 11, 761-797.

[11] Carr P., Madan D. 1999. Option valuation using the fast Fourier transform. *Journal of Computational Finance* 2 (4): 61–73.

[12] Maheu J., Chan W. 2002. Conditional Jump Dynamics in Stock Market Returns, *Journal of Business & Economic Statistics*, 20(3), 377-389

[13] Carr P., Wu L. 2016. Leverage Effect, Volatility Feedback, and Self-Exciting Market Disruptions. *Forthcoming in Journal of Financial and Quantitative Analysis.*

[14] Chavez-Demoulin, V., Davison, A.C., McNeil, A.J. 2005. A point process approach to value-at-risk estimation. *Quantitative Finance*, 5 (2), 227–234.

[15] Chavez-Demoulin, V., McGill J. 2012. High-frequency financial data modeling using Hawkes processes. *Journal of Banking and Finance*. 36, 3415–3426.

[16] Chen K., Poon S.-H. 2013. Variance swap premium under stochastic volatility and self-exciting jumps. *working paper* SSRN-id2200172.

[17] Chernov, Mikhail, A. Ronald Gallant, Eric Ghysels and George Tauchen 1999. A New Class of Stochastic Volatility Models with Jumps: Theory and Estimation, working paper, Cirano.

[18] Da Fonseca, J., Zaatour, R. 2014. Hawkes Process: Fast Calibration, Application to Trade Clustering, and Diffusive Limit. *Journal of Futures Markets*. 34(6), p548-579

[19] Demeterfi K., Derman E., Kamal M., and Zou J. More than you ever wanted to know about volatility swaps. *Quantitative Strategies Research Notes*, Goldman Sachs 1999.

[20] Doucet, A., J.F.G. De Freitas and N. Gordon. *Sequential Monte Carlo Methods in Practice*. Cambridge University Press, Cambridge 2000.

[21] Duffie D., Pan J., Singleton K. 2000. Transform analysis and asset pricing for affine jump diffusions, *Econometrica* 68, 1343-1376.

[22] Embrechts, P., Liniger, T., Lu, L. 2011, Multivariate Hawkes processes: an application to financial data. *Journal of Applied Probability*, 48 (A), 367–378.

[23] Eglo, D., Leippold, M., Wu, L. 2010. The term structure of variance swap rates and optimal variance swap investments. *Journal of Finance and Quantitative Analysis*, 45, 1279-1310.

[24] Engle, R. F., Russell, J. R., 1998. Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data, *Econometrica*, 66, 1127-1162

[25] Engle R.F. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50, 987-1007.

[26] Eraker, B., 2004. Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *Journal of Finance* 59, 1367-1403.

[27] Eraker, B., Johannes, M., Polson, N., 2003. The impact of jumps in volatility and returns. *Journal of Finance*. 58, 1269-1300.

[28] Errais, E., Giesecke, K., Goldberg, L.R. 2010. Affine point processes and portfolio credit risk. *SIAM Journal of Financial Mathematics*. 1, 642-665.

[29] Evans K. 2011. Intraday jumps and US macroeconomic news announcements. *Journal of Banking & Finance*, 35 (10), 2511-2527.

[30] Fulop A., Li J., Yu J. 2015. Self-Exciting Jumps, Learning, and Asset Pricing Implications. *Review of Financial Studies* 28, 876-912.

[31] Fortune P., 1999. Are stock returns different over weekends? a jump diffusion analysis of the "weekend effect". *New England Economic Review*.3-19

[32] Gallant, A. Ronald, and George Tauchen, 1996, Which moments to match? *Econometric Theory* 12, 657–681.

[33] Gourieroux C., Monfort A., Renault E., 1993, Indirect inference, *Journal of Applied Econometrics* 8, S85–S118

[34] Giot, P. 2005. Market risk models for intraday data. *European Journal of Finance*. 11 (4), 309–324.

[35] Hainaut D. 2016. A model for interest rates with clustering effects. *Quantitative Finance*, forthcoming.

[36] Heston S.L. 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6, 327-343.

[37] Hawkes, A., 1971(a). Point sprectra of some mutually exciting point processes. *Journal of the Royal Statistical Society Series B*, 33, 438-443.

[38] Hawkes, A., 1971(b). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58, 83–90.

[39] Hawkes A. and Oakes D., 1974. A cluster representation of a self-exciting process. *Journal of Applied Probability*, 11, 493-503.

[40] Hewlett P. 2006. Clustering of order arrivals, price impact and trade path optimisation. *Workshop on Financial Modeling with Jump processes*, Ecole Polytechnique.

[41] Kou S.G. 2002. A Jump-Diffusion Model for Option Pricing.*Management Science*, 48(8), 1086-1101.

[42] Large J. 2007. Measuring the resiliency of an electronic limit order book, *Journal of Financial Markets* 10(1), 1-25

[43] Lee S. 2012. Jumps and Information Flow in Financial Markets. *Review of financial studies* 25 (2), 439-479.

[44] Lee S. Mykland P. 2008. Jumps in financial markets: a new nonparametrix test and jump dynamics. *Review of Financial Studies* 21 (6), 2535-2563.

[45] Maheu, J., McCurdy, T., 2004. News arrival, jump dynamics and volatility components for individual stock returns. *Journal of Finance* 59, 755-793

[46] Merton, R. 1976. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* 3, 125-144.

[47] Ncube M.M. Stochastic models and inferences for commodity futures pricing. PhD Thesis Florida State University, 2009.

[48] Pan, J., 2002. The jump-risk premia implicit in options: evidence from an integrated time-series study. *Journal of Financial Economics* 63, 3-50.

[49] Press. J. 1967. A compound events model for security prices. *Journal of Business* 40. 317-335.

[50] Rangel J. 2011. Macroeconomic news, announcements, and stock market jump intensity dynamics. *Journal of Banking and Finance* 35 (5), 1263-1276.

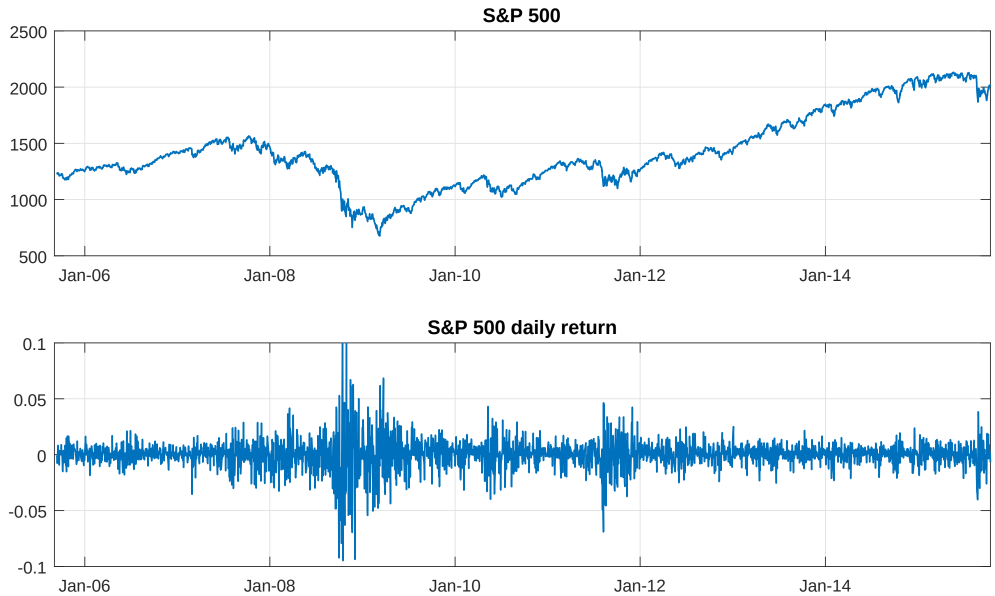
[51] Schoutens W. 2003. *Lévy processes in Finance. Pricing of financial derivatives.* Eds Wiley.

[52] Yu J. 2004. Empirical characteristic function estimation and its applications. *Econometric reviews* 23 (2) p 93-123.

[53] Zhang X.W. , Glynn P.W., Giesecke K., Blanchet J. 2009. Rare event simulation for a generalized Hawkes process. in *Proceedings of the 2009 Winter Simulation Conference.*

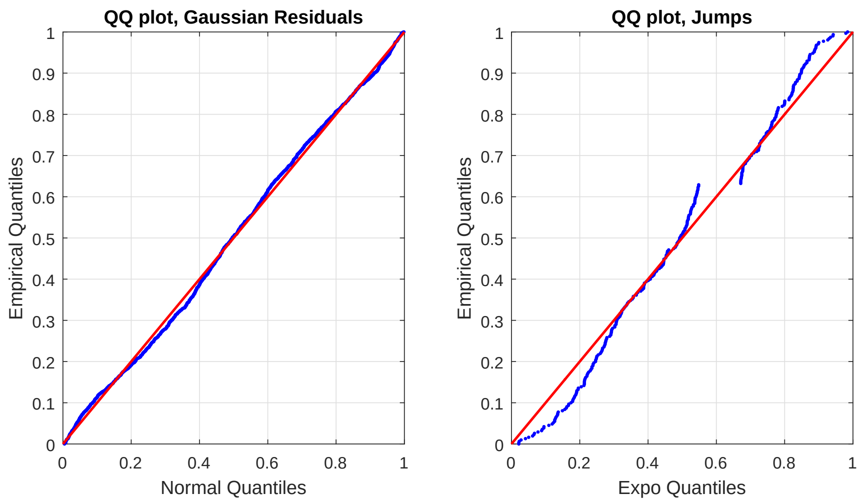
7 Figures

Figure 1: Prices and returns of the S&P 500



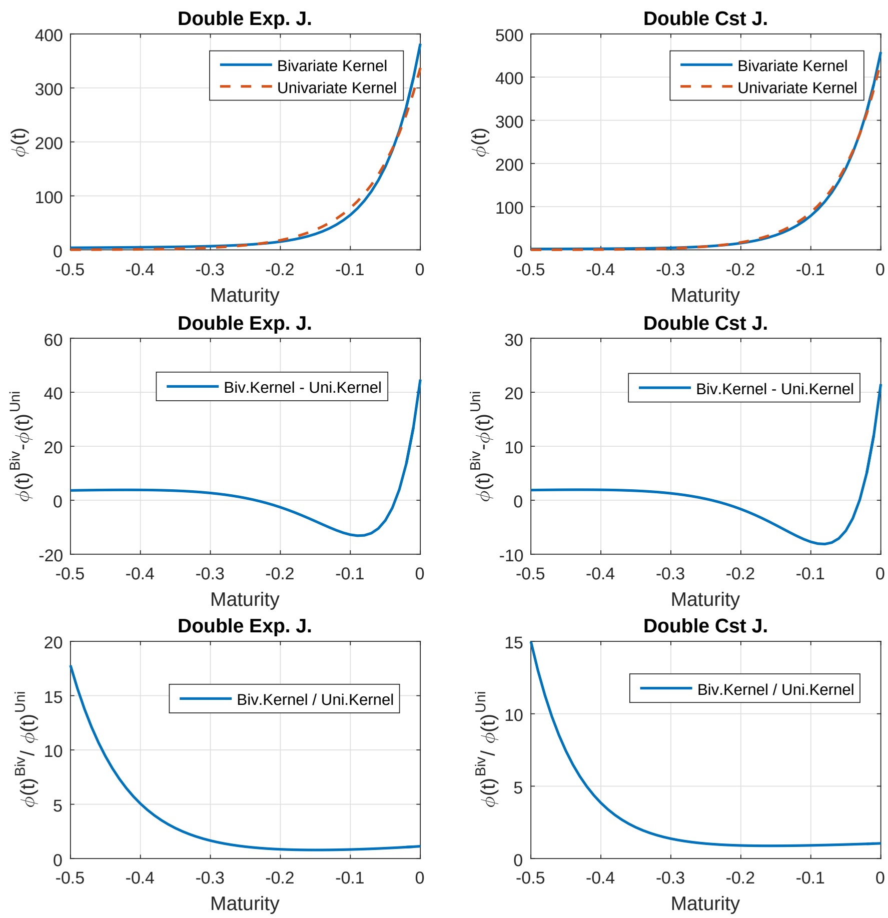
This figure presents the S&P 500 daily quotes and log returns from the 7/9/2005 to the 13/10/2015.

Figure 2: QQ plots



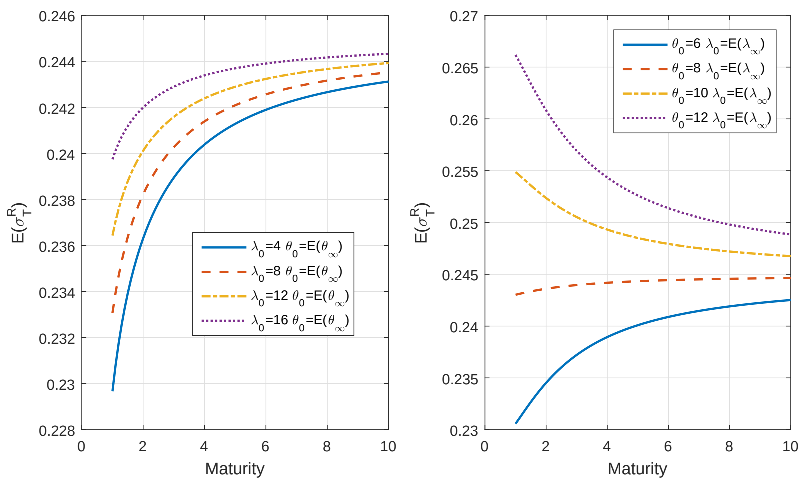
First and second graphs: QQ plots of filtered Gaussian residuals and double exponential jumps. Notice that filtered jumps smaller than 0.4% have been excluded from the sample to draw the QQ plot because they form a white noise generated by the particles filter.

Figure 3: Kernel functions



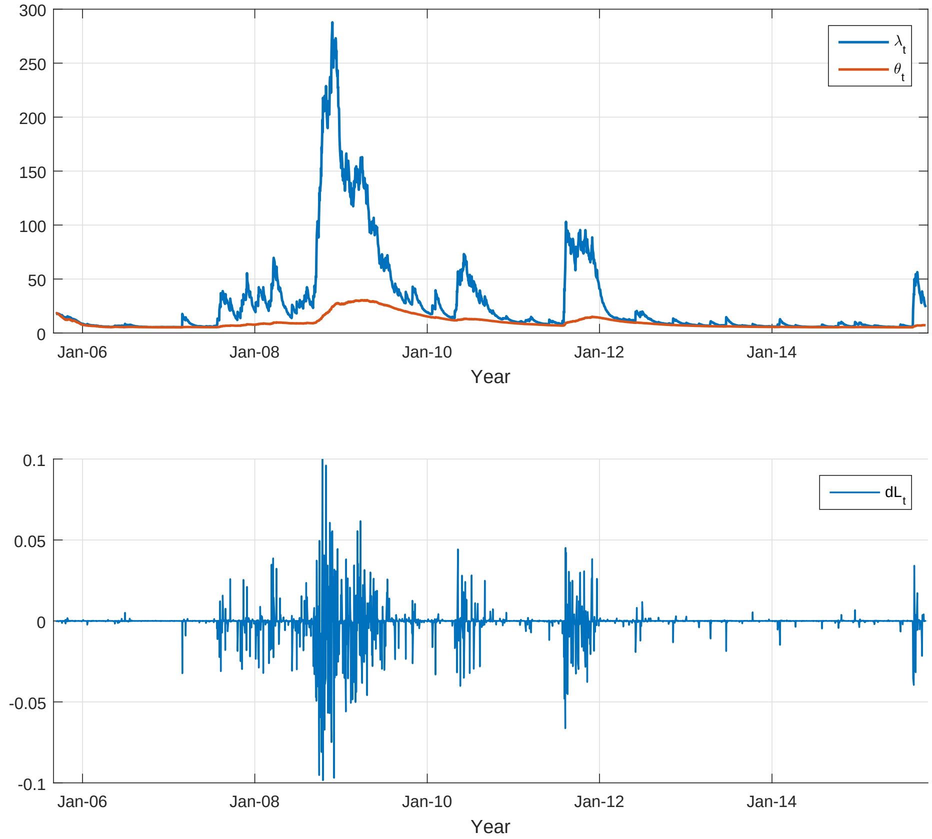
Right and left graphs compare kernel functions of uni- and bi- factors model, respectively for double constant and double exponential jumps. The first, second and third rows show respectively kernel functions, their differences and their ratio.

Figure 4: Expected realized volatilities



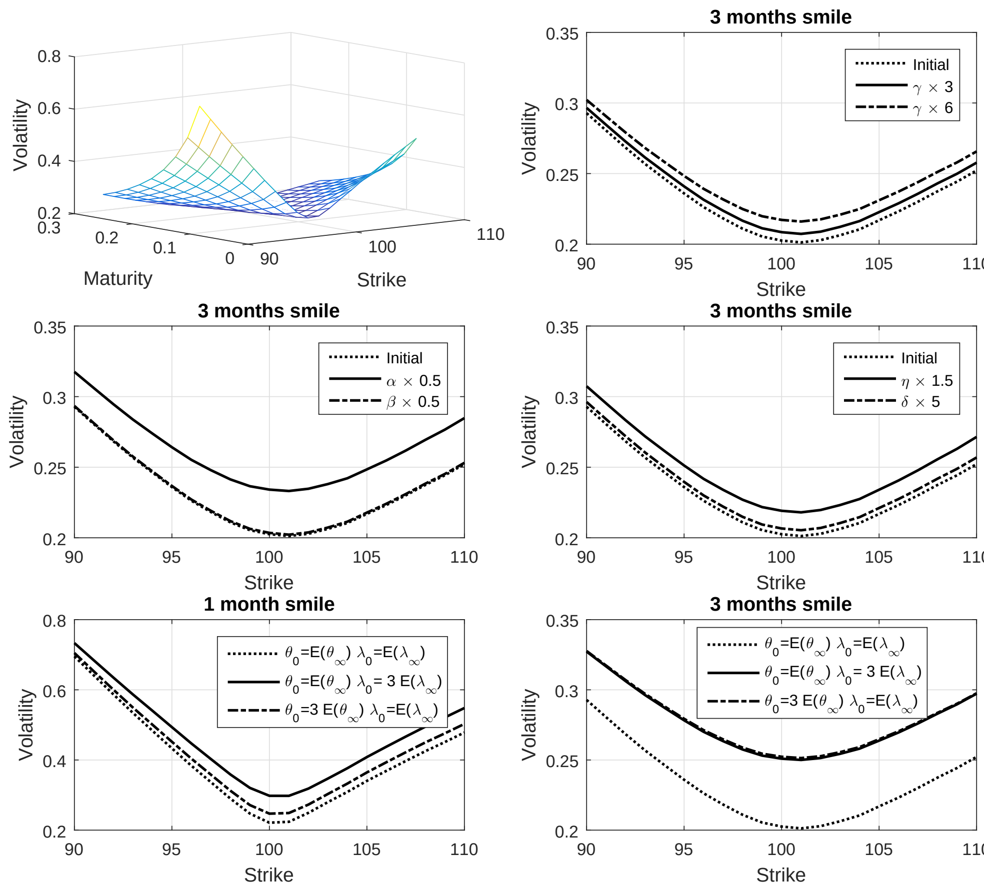
Curves of expected realized volatilities for different values of λ_0 and θ_0.

Figure 5: Filtered processes



Filtered values of λ_t, θ_t and jumps, for the DEJ model.

Figure 6: Surface and smiles of implied volatilities



Left upper graph: surface of implied volatilities for European call options on the S&P 500. The other graphs analyse the sensitivity of the volatility smile to change of parameters and of initial value λ_0 and θ_0.

8 Tables

Table 1: Descriptive Statistics for Returns Data

	Value
Mean daily return	0.02%
Standard daily deviation	1.30%
Skewness	-0.33
Kurtosis	13.51
Jarque Bera p-value	1e-3
Lillie test p-value	1e-3
Durbin Watson p-value	0e-3

This table reports the mean, the standard deviation, the skewness, the kurtosis, statistics of normality and serial dependence, for the continuously compounded returns of the S&P 500 from September 2005 to October 2015.

Table 2: Filtered loglikelihood

	Bi factor model			
	D.Exp. J.	Pos. Exp. J.	Neg. Exp. J.	Bin. J.
Log. Lik.	6999	6193	6388	6950
AIC	-13979	-12367	-12756	-13879
	Uni factor model			
	D.Exp. J.	Pos. Exp. J.	Neg. Exp. J.	Bin. J.
Log. Lik.	6989	6193	6384	6944
AIC	-13962	-12371	-12752	-13872
LRT	20	0	8	12
p-value	0.00%	100%	1.83%	0.25%

This table reports the log-likelihood, AIC filtered by the particle filter for the double and single factor models and the p-value of the log-likelihood ratio test.

Table 3: Parameters Estimation

Bi factor model				
	D.Exp. J.	Pos. Exp. J.	Neg. Exp. J.	Bin. J.
μ	0.05	0.05	0.05	0.05
σ	0.12	0.18	0.16	0.12
α	18.78	13.19	9.91	17.95
η	381.80	291.40	220.17	458.06
β	1.77	1.62	1.36	1.44
γ	5.07	1.71	4.28	4.07
δ	8.37	4.68	3.80	4.07
p	0.37	-	-	0.37
ρ^+	30.47	30.47	-	30.47
ρ^-	-33.90	-	-33.90	-33.90

Uni factor model				
	D.Exp. J.	Pos. Exp. J.	Neg. Exp. J.	Bin. J.
μ	0.05	0.05	0.05	0.05
σ	0.12	0.18	0.16	0.128
α	14.71	11.46	8.73	16.17
η	337.08	273.32	208.82	436.55
θ	6.44	2.05	4.78	4.88
p	0.37	-	-	0.37
ρ^+	30.47	30.47	-	30.47
ρ^-	-33.90	-	-33.90	-33.90

This table reports the parameters of the double and single factor models fitted with the POT procedure.

Table 4: Asymptotic moments

	Bi factor model			
	D.Exp. J.	Pos. Exp. J.	Neg. Exp. J.	Bin. J.
$\mathbb{E}(\theta_\infty)$	8.27	2.61	5.62	6.41
$\mathbb{E}(\lambda_\infty)$	22.03	9.46	16.30	29.69
$\sqrt{\mathbb{V}(\lambda_\infty)}$	13.01	8.22	8.46	12.92
	Uni factor model			
	D.Exp. J.	Pos. Exp. J.	Neg. Exp. J.	Bin. J.
$\mathbb{E}(\lambda_\infty)$	21.80	9.42	16.18	28.61
$\sqrt{\mathbb{V}(\lambda_\infty)}$	12.63	8.13	8.38	12.63

This table reports the the asymptotic values of intensities $\mathbb{E}(\theta_\infty)$, $\mathbb{E}(\lambda_\infty)$ and their standard deviations.