

DISCUSSION PAPER

2018/13

A switching self-exciting jump diffusion
process for stock prices

Hainaut, H. and F. Moraux

A switching self-exciting jump diffusion process for stock prices

Donatien Hainaut*

ISBA, Université Catholique de Louvain, Belgium

Franck Moraux†

Univ. Rennes 1 and CREM

May 25, 2018

Abstract

This study proposes a new Markov switching process with clustering effects. In this approach, a hidden Markov chain with a finite number of states modulates the parameters of a self-excited jump process combined to a geometric Brownian motion. Each regime corresponds to a particular economic cycle determining the expected return, the diffusion coefficient and the long-run frequency of clustered jumps. We study first the theoretical properties of this process and we propose a sequential Monte-Carlo method to filter the hidden state variables. We next develop a Markov Chain Monte-Carlo procedure to fit the model to the S&P 500. Finally, we analyse the impact of such a jump clustering on implied volatilities of European options.

KEYWORDS : switching regime, Hawkes process, self-excited jumps

MSC Codes: 60J60, 60J75

1 Introduction.

There is considerable empirical evidence suggesting that the random walk model for changes in stock prices is not appropriate due to the clustering of jumps. Work by Osborne (1963), Alexander (1964), and Mandelbrot (1963) find that transactions in a given stock are not independent of the history of trades in that stock. This also explains the success of the ARCH (Engel, (1982)) and GARCH (Bollerslev (1986)) models for stock prices. These models set conditional variance equal to a constant plus a weighted average of past squared residuals. Jump-diffusion models and Lévy processes offer an interesting alternative in continuous time. These processes can accommodate the skewness and the excess kurtosis that we observe in financial security returns¹. However their increments are independent and identically distributed (see e.g. Schoutens, 2003) and their dynamics cannot exhibit volatility and jump clustering of stocks returns.

On the other hand, the clustering of jumps has been the focus of much research and investigation. Jump clustering is a characteristic of any jump model equipped with a time-varying and persistent conditional intensity (see Maheu and Chan (2002), for a typical example). 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 stochastic volatility. Duffie et al. (2000) generalize the approach by assuming that the jump intensity is an affine function of a latent variable.

*Postal address: Voie du Roman Pays 20, 1348 Louvain-la-Neuve (Belgium) . E-mail to: donatien.hainaut(at)uclouvain.be

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

¹Notice that stochastic volatility models also explain asymmetry and high kurtosis. However Cont and Tankov (2004) mention on page 6 that Brownian models with stochastic volatility cannot explain the presence of jumps in price due to the continuity of sample paths.

This jump clustering phenomenon questions the traditional understanding of equity market dynamics and the tools commonly used by practitioners and academics. A recent strand of the literature makes significant efforts to understand the cause of the clustering of jumps and to investigate implications of it for asset pricing, option pricing and risk management. The reason why we observe jumps in equity prices has been especially investigated with high-frequency data and there is a broad consensus in this literature that news releases impact at both the individual and market levels (see e.g. Lee and Mykland (2008) and Evans (2011) for additional evidence). More generally, it is widely believed that the information flow matters and as noticed by Maheu and McCurdy (2004) and that this flow explain some clustering effects.

A recent endogenous way to model the clustering of jumps is provided by self-exciting point processes. In this category of processes, the jump arrival intensity at a given point in time (that is also the instantaneous probability to observe a jump) depends on the number of past jumps of price. This approach is linked to the Hawkes self-exciting processes (see Hawkes (1971a, b) and Hawkes and Oakes (1974)). In the most common and simplest specification, the jump arrival intensity process is persistent and it suddenly increases when a jump occurs in the asset price. Moreover, the influence of a jump 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 excitation in order to question if some jumps are caused by contagion effects. Carr and Wu (2016), Chen and Poon (2013) and Fulop et al. (2015) rather develop mono-asset settings in which only negative jumps are self-excited. Hainaut (2016) develops a model for interest rates based on Hawkes processes.

Existing self-excited processes are partly unsatisfactory because they fail to account for economic cycles. In particular, we will show in this article that self-excited shocks occur mainly during economic recessions and nearly disappear in periods of economic growth. Ignoring this aspect leads to underestimate the mean level toward which the intensity of jumps reverts during financial turmoil and to overestimate the contagion risk during economic recovery. In order to demonstrate this phenomenon, we develop a model in which a self-excited jump diffusion process is modulated by a hidden Markov chain. This model is called the “switching self-excited jump diffusion” (SSEJD). In this approach, each state of the Markov chain corresponds to an economic cycle that drives the behaviour of assets and the clustering of shocks. According to our knowledge, the literature on this approach is almost nonexistent. Wang et al. (2012) explore Markov-modulated Hawkes processes with stepwise decay. However, their framework significantly differs from our model. Firstly, the intensity of jumps switches between several Hawkes processes at discrete occurrence times, and are constant between these times for computational feasibility. Secondly, the Hawkes process has constant increments and is not combined with a Brownian motion. Thirdly, as the Hawkes process is hidden in our model, we need to develop a particle Monte Carlo Markov Chain method for the calibration. Elliott et al. (2007) propose a switching jump diffusion process for option pricing where jumps are not self-excited. Siu (2016) study a self-exciting threshold jump–diffusion model in which regime switches in the price process are introduced via a set of threshold parameters over the state space of the previous price. Consequently, regime switches are incorporated without introducing an exogenous stochastic factor process.

Economic cycles and their modeling by a hidden Markov chain received a lot of attention in the econometric literature. Our article is directly related to this line of research. For example, Guidolin and Timmermann (2005) present evidence of persistent ‘bull’ and ‘bear’ regimes in UK stock and bond returns and considers their economic implications from the perspective of an investor’s portfolio allocation. Similar results are found in Guidolin and Timmermann (2008), for international stock markets. Guidolin and Timmermann (2007) develop a regime switching model for asset returns with four states. From a broader perspective, regime switching models are useful to describe equity markets. Cholette et al. (2009) fit skewed-t GARCH marginal distributions for international equity returns and a regime switching copula with two states. Al-Anaswah and Wilfing (2011) propose a two-regime Markov-switching specification to

capture speculative bubbles. On the other hand, Calvet and Fisher (2001,2004) show that discrete versions of multi fractal processes capture thick tails and have a switching regime structure. Nakajima (2013) proposes a bayesian analysis of the stochastic volatility model with regime-switching skewness. Switching regime processes are also used by practitioners for investment management purposes (see Hunt and Kavesh (1976), Hunt (1987) or Stovall (1996)). Finally, Hardy (2001) and the society of actuaries (SOA) since 2004, recommend switching processes to model long term stocks return, in actuarial applications.

The paper proceeds as follows. Section 2 introduces the switching self-excited jump diffusion (SSEJD) and its features. We infer the first and second moments of the jump intensity and the moment generating function of log-returns. Section 3 focuses on the estimation of parameters. We develop first a particle filter to estimate hidden state variables. We next present a peak over threshold (POT) method to fit the process. And finally, we propose a particle Monte Carlo Markov Chain (PMCMC) algorithm that adjusts the parameters found with the POT procedure. The model is fitted to the S&P 500 time serie of daily returns. Section 4 validates our estimation approach. Section 5 presents a numerical framework for options pricing with the SSEJD. We conclude by a analysis of sensitivity of option prices.

2 The SSJED model

2.1 The framework

We consider a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with an augmented filtration $\mathbb{F} := \{\mathcal{F}_t\}_{t \geq 0}$ that is generated by a Markov chain and another processes to be specified below. Here \mathbb{P} denotes the real-world probability measure. Throughout this paper, the economy is categorized into N states or regimes, indexed by a set of integers $\mathcal{N} := \{1, 2, \dots, N\}$. The information about the economic state over time, is carried by a hidden Markov chain $\delta(t)$ taking values from a set of unit vectors $E = \{e_1, \dots, e_N\}$, where $e_j = (0, \dots, 1, \dots, 0)'$ (null vector except the j^{th} element that is equal to 1). The natural filtration generated by $\{\delta(t)\}_{t \geq 0}$ is denoted by $\{\mathcal{G}_t\}_{t \geq 0}$. The generator of $\delta(t)$ is an $N \times N$ matrix $Q_0 := [q_{i,j}]_{i,j=1,2,\dots,N}$, whose elements satisfy the following standard conditions:

$$q_{i,j} \geq 0, \quad \forall i \neq j, \quad \text{and} \quad \sum_{j=1}^N q_{i,j} = 0, \quad \forall i \in \mathcal{N}. \quad (1)$$

$q_{i,j}\Delta$ is the probability that the Markov chain transits from e_i to e_j , with $i \neq j$ over a short period of time Δ . The matrix of transition probabilities over the time interval $[t, s]$ is denoted as $P(t, s)$ and is the matrix exponential of the generator matrix, times the length of the time interval:

$$P(t, s) = \exp(Q_0(s - t)), \quad s \geq t. \quad (2)$$

The elements of this matrix, $p_{i,j}(t, s)$, $i, j \in \mathcal{N}$, defined as

$$p_{i,j}(t, s) = P(\delta(s) = e_j \mid \delta(t) = e_i), \quad i, j \in \mathcal{N}, \quad (3)$$

are the probabilities of switching from state i at time t to state j at time s . The probability of the chain being in state i at time t , denoted by $p_i(t)$, depends upon the initial probabilities $p_k(0)$ at time $t = 0$ and the transition probabilities $p_{k,i}(0, t)$, where $k = 1, 2, \dots, N$, as follows:

$$p_i(t) = P(\delta(t) = e_i) = \sum_{k=1}^N p_k(0) p_{k,i}(0, t), \quad \forall i \in \mathcal{N}. \quad (4)$$

Elliott et al. (2005) provide a semi-martingale representation theorem for $\delta(t)$. If we remember that \mathcal{G}_t is the filtration of $\delta(t)$, we can show that $\delta(t)$ is the sum of an \mathcal{G}_t predictable-adapted process and of a \mathcal{G}_t

martingale increment process, $\{M_t\}_{t \geq 0}$:

$$\delta(t) = \delta(0) + \int_0^t Q'_0 \delta(s) ds + M_t. \quad (5)$$

We define the price process S_t of a financial asset. The instantaneous return of this asset is the sum of a drift, a Brownian motion W_t , and a compensated jump process:

$$\begin{aligned} \frac{dS_t}{S_{t-}} &= \mu_t dt + \sigma_t dW_t + d \left(\sum_{j=1}^{N_t} (e^{J_j} - 1) \right) - \lambda_t \mathbb{E} (e^J - 1) dt \\ &= \mu_t dt + \sigma_t dW_t + (e^J - 1) dN_t - \lambda_t \mathbb{E} (e^J - 1) dt \end{aligned} \quad (6)$$

The drift rate μ_t , and the Brownian volatility σ_t , are modulated by the Markov chain δ . That is, $\mu_t = \delta(t)' \bar{\mu}$ and $\sigma_t = \delta(t)' \bar{\sigma}$ where $\bar{\mu} = (\bar{\mu}_1, \dots, \bar{\mu}_N)' \in \mathbb{R}^N$ and $\bar{\sigma} = (\bar{\sigma}_1, \dots, \bar{\sigma}_N)' \in \mathbb{R}_+^N$. $(\lambda_t)_{t \geq 0}$ is the intensity of the counting process $(N_t)_{t \geq 0}$. The natural filtration of (S_t, λ_t, N_t) is denoted by $(\mathcal{H}_t)_{t \geq 0}$. The filtration $(\mathcal{F}_t)_{t \geq 0}$ is the augmented filtration $\mathcal{F}_t = \mathcal{H}_t \vee \mathcal{G}_t$ that carries the information about the economic regime, the jump process and the asset price.

$(N_t)_{t \geq 0}$ is a point process and J_j are i.i.d. random jumps with a density denoted by $\nu(z)$ on \mathbb{R} . The last term of the equation (6) is the compensator of the jump process. The presence of the compensator ensures that the expected return is equal to $\mathbb{E} \left(\frac{dS_t}{S_{t-}} \right) = \mu_t dt$. Applying the Itô's lemma allows us to rewrite the dynamics of $\ln S_t$ as follows

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

and we infer by direct integration of this relation, the next expression for S_t :

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

We assume that jumps are double-exponential random variables (DEJ). However, most of results developed in this work are applicable to any other distributions. The probability distribution function (pdf) of DEJ, noted here $\nu(z)$, is defined by the three parameters $\rho^+ \in \mathbb{R}^+$, $\rho^- \in \mathbb{R}^-$ and $p \in (0, 1)$:

$$\nu(z) = p \rho^+ e^{-\rho^+ z} 1_{\{z \geq 0\}} - (1-p) \rho^- e^{-\rho^- z} 1_{\{z < 0\}}. \quad (7)$$

where p and $(1-p)$ are respectively the probabilities of observing upward and downward exponential jumps. Average sizes of positive and negative shocks are equal to $\frac{1}{\rho^+}$ and $\frac{1}{\rho^-}$. The expectation of J_j is the weighted sum of expected average jumps:

$$\mathbb{E}(J_j) = p \frac{1}{\rho^+} + (1-p) \frac{1}{\rho^-}. \quad (8)$$

In later developments, the moment-generating function for the sum of J and of its absolute value is needed. We define it as follows:

$$\begin{aligned} \psi(z_1, z_2) &:= \mathbb{E} \left(e^{z_1 J + z_2 |J|} \right) \\ &:= p \frac{\rho^+}{\rho^+ - (z_1 + z_2)} + (1-p) \frac{\rho^-}{\rho^- - (z_1 - z_2)} \end{aligned} \quad (9)$$

under the condition that $(z_1 + z_2) < \rho^+$ and $(z_1 - z_2) > \rho^-$. Notice that

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

When λ_t is constant, the model corresponds to a switching regime version of the Double Exponential Jump Diffusion (DEJD) proposed by Kou (2002). As underlined in the introduction, this model fails to duplicate the clustering of shocks observed in stocks markets, as its moments are constant. In order to introduce such a feature in the DEJD, we assume that the frequency of jumps arrival λ_t is itself a stochastic process, that depends upon the sum of absolute values of past jumps. This sum up to time t is denoted by L_t and is equal to

$$L_t := \sum_{i=1}^{N_t} |J_i|. \quad (10)$$

The frequency λ_t reverts to a level θ_t , at a speed α and increases of $\eta|J|$ ($\eta \in \mathbb{R}^+$) when a jump occurs:

$$d\lambda_t = \alpha(\theta_t - \lambda_t)dt + \eta dL_t. \quad (11)$$

The level of mean reversion θ_t is modulated by the Markov chain: $\theta_t = \delta(t)' \bar{\theta}$ where $\bar{\theta} = (\bar{\theta}_1, \dots, \bar{\theta}_N)' \in \mathbb{R}_+^N$. This specification allows us to differentiate the clustering effects feature between the different economic regimes. In a state corresponding to a period of economic turmoil, shocks are more frequent and grouped. We expect then a high mean reversion level. At the contrary, jumps are rare in period of economic growth and the mean reversion level in this state tends to be low. The couple (λ_t, θ_t) is a Markov process. We can show that

$$\lambda_t = \lambda_0 - \alpha \int_0^t e^{\alpha(s-t)} (\lambda_0 - \theta_s) ds + \int_0^t \eta e^{\alpha(s-t)} dL_s. \quad (12)$$

To prove this relation, we differentiate the expression of λ_t to retrieve the equation (11):

$$\begin{aligned} d\lambda_t &= -\alpha(\lambda_0 - \theta_t)dt + \eta dL_t + \alpha^2 \int_0^t e^{\alpha(s-t)} (\lambda_0 - \theta_s) ds dt - \alpha \int_0^t \eta e^{\alpha(s-t)} dL_s dt. \\ &= \alpha\theta_t - \alpha \left(\lambda_0 - \alpha \int_0^t e^{\alpha(s-t)} (\lambda_0 - \theta_s) ds + \int_0^t \eta e^{\alpha(s-t)} dL_s \right) + \eta dL_t \\ &= \alpha(\theta_t - \lambda_t)dt + \eta dL_t \end{aligned}$$

Notice that α , η and parameters of J are invariant with respect to the economic state. There are two reasons for this choice. The first one is our willingness to preserve a certain parsimony. Secondly if they depend on $\delta(t)$, we lose the analytical tractability as it is no more possible in this case to establish a relation similar to equation (12). α being constant, markets recover from economic shocks at the same speed whatever the conjuncture. A constant η means that the relative sensitivity of investors to a shock remains the same over time. However, in absolute terms the sensitivity of investors to shocks, that is ηdL_t , depends upon the economic regime through the number of jumps N_t .

2.2 Features of the jump component

This section provides various results and analytical expressions that help to understand the behaviour of the jump component in the SSEJD process. Most of results are based on the Markov property of the triplet $(\lambda_t, \delta(t), N_t)$. The first proposition is about the expected intensities of jumps.

Proposition 2.1. *The expected value of λ_t , conditionally to \mathcal{F}_0 , is given by the next expression:*

$$\begin{aligned} \mathbb{E}(\lambda_t | \mathcal{F}_0) &= \alpha \delta(0)' (Q_0 - I(\eta \mathbb{E}(|J|) - \alpha))^{-1} \\ &\quad \times [\exp(Q_0 t) - \exp(I(\eta \mathbb{E}(|J|) - \alpha)t)] \bar{\theta} + \lambda_0 e^{(\eta \mathbb{E}(|J|) - \alpha)t} \end{aligned} \quad (13)$$

where I is the identity matrix of size $N \times N$.

Proofs of propositions presented in this article are detailed in appendix A. An important consequence of this proposition is that the process is stable (in the sense that its limit exists when $t \rightarrow \infty$) only if $\eta \mathbb{E}(|J|) - \alpha \leq 0$. As a result, the speed of mean reversion is constrained to be larger than $\eta \mathbb{E}(|J|)$. In this case, the asymptotic value to which $\mathbb{E}(\lambda_t | \mathcal{F}_0)$ converges when t tends to infinity is finite and equal to

$$\lambda_\infty := \lim_{t \rightarrow \infty} \mathbb{E}(\lambda_t | \mathcal{F}_0) = \alpha \delta(0)' (Q_0 - I(\eta \mathbb{E}(|J|) - \alpha))^{-1} P_\infty \bar{\theta} \quad (14)$$

where P_∞ is the matrix of steady state probabilities that is such that $\delta(0)' P_\infty = \pi$. Remark that the stability condition is independent from the economic regime. The next corollary is useful to derive the ordinary differential equation, satisfied by the second moment of λ_t :

Corollary 2.2. *Denote by $\bar{\theta}^2$ the N vector of $(\theta_j^2)_{j=1, \dots, N}$, the expected value of $\theta_t \lambda_t$, conditionally to \mathcal{F}_0 , is :*

$$\begin{aligned} \mathbb{E}(\theta_t \lambda_t | \mathcal{F}_0) &= \theta_0 \lambda_0 e^{(\eta \mathbb{E}(|J|) - \alpha)t} + \alpha \delta(0)' (Q_0 - I(\eta \mathbb{E}(|J|) - \alpha))^{-1} \\ &\quad \times [\exp(Q_0 t) - \exp(I(\eta \mathbb{E}(|J|) - \alpha)t)] \bar{\theta}^2. \end{aligned} \quad (15)$$

where I is the identity matrix of size $N \times N$.

Proposition 2.3. *The second moment of λ_t is solution of the following ordinary differential equation:*

$$\begin{aligned} \frac{\partial}{\partial t} \mathbb{E}(\lambda_t^2 | \mathcal{F}_0) &= 2\alpha \mathbb{E}(\theta_t \lambda_t | \mathcal{F}_0) + 2(\eta \mathbb{E}(|J|) - \alpha) \mathbb{E}(\lambda_t^2 | \mathcal{F}_0) \\ &\quad + \eta^2 \mathbb{E}(\lambda_t | \mathcal{F}_0) \mathbb{E}(|J|^2) \end{aligned} \quad (16)$$

with the initial condition: $\mathbb{E}(\lambda_t^2 | \mathcal{F}_0) = \lambda_0^2$. $\mathbb{E}(\theta_t \lambda_t | \mathcal{F}_0)$ and $\mathbb{E}(\lambda_t | \mathcal{F}_0)$ are respectively provided by expressions (15) and (13).

With these propositions, we calculate the variance of λ_t . The next proposition presents the probability generating function of the number of jumps which is the exponential of an affine function of the intensity:

Proposition 2.4. *The probability generating function of N_s with $s \geq t$ is given by*

$$\mathbb{E}(\omega^{N_s} | \mathcal{F}_t) = \omega^{N_t} \exp(A(t, s, \delta(t)) + B(t, s) \lambda_t) \quad .$$

where $B(t, s)$ is the solution of an ODE:

$$\frac{\partial}{\partial t} B(t, s) = \alpha B - [\omega \psi(0, B(t, s) \eta) - 1] \quad (17)$$

with the terminal condition $B(s, s) = 0$ and where $\tilde{A}(t, s) = [e^{A(t, s, e_1)}, \dots, e^{A(t, s, e_N)}]^\top$ is a vector, solution of the ODE system:

$$\frac{\partial \tilde{A}(t, s)}{\partial t} + (\text{diag}(\alpha \bar{\theta} B) + Q_0) \tilde{A}(t, s) = 0, \quad (18)$$

under the terminal boundary condition:

$$\tilde{A}(T, j) = 1 \quad j = 1 \dots N.$$

2.3 The behaviour of log-returns

This section provides results about the log-return of the financial asset in the SSEJD model. Throughout the paper, the log return of S_t is denoted by $X_t := \ln \frac{S_t}{S_0}$. The dynamic of X_t that obeys to the next SDE:

$$dX_t = \left(\mu_t - \frac{\sigma_t^2}{2} - \lambda_t \mathbb{E}(e^J - 1) \right) dt + \sigma_t dW_t + J dN_t.$$

The next proposition presents the moment generating function (mgf) of X_t . This mgf is used in the last section to evaluate options by a Fast Fourier Transform. The mgf can also be inverted numerically to retrieve the joint distribution of X_t and λ_t .

Proposition 2.5. *The mgf of $\omega_1 X_s + \omega_2 \lambda_s$ for $s \geq t$, is given by the following expression*

$$\mathbb{E} \left(e^{\omega_1 X_s + \omega_2 \lambda_s} \mid \mathcal{F}_t \right) = \left(\frac{S_t}{S_0} \right)^{\omega_1} \exp(A(t, s, \delta(t)) + B(t, s) \lambda_t),$$

where $B(t, s)$ is a function of time, solution of the ODE:

$$\frac{\partial}{\partial t} B(t, s) = \alpha B + \omega_1 (\psi(1, 0) - 1) - [\psi(\omega_1, B \eta) - 1] \quad (19)$$

with the terminal condition $B(s, s) = \omega_2$ and where $\tilde{A}(t, s) = [e^{A(t, s, e_1)}, \dots, e^{A(t, s, e_N)}]^\top$ is a vector of functions, solution of the ODE system:

$$\frac{\partial \tilde{A}(t, s)}{\partial t} + \left(\text{diag} \left(\omega_1 \left(\bar{\mu} - \frac{\bar{\sigma}^2}{2} \right) + \omega_1^2 \frac{\bar{\sigma}^2}{2} + \alpha \bar{\theta} B \right) + Q_0 \right) \tilde{A}(t, s) = 0, \quad (20)$$

under the terminal boundary condition:

$$\tilde{A}(s, s) = 1 \quad j = 1 \dots N.$$

At our knowledge, there does not exist any closed form expression for $B(t, s)$ or $A(t, s, \delta(t))$. However the ODEs defining then may be solved numerically (e.g. by an Euler method). On the other hand, $B(t, s)$ admits the following integral representation:

Proposition 2.6. *If we note*

$$\beta(\omega_1) = \omega_1 (\psi(1, 0) - 1) + 1$$

then $B(t, s) = F_{\omega_1}^{-1}(s - t)$ where

$$F_{\omega_1}(C) := \int_{\omega_2}^C \frac{du}{-\beta(\omega_1) - \alpha u + \psi(\omega_1, \eta u)},$$

and $C \in [\omega_2, u^*)$ or $C \in [u^*, \omega_2)$ where u^* is the unique solution of ²

$$\psi(\omega_1, u\eta) = \beta + \alpha u.$$

²Notice that if $C \in [u^*, \omega_2)$, the function $F_{\omega_1}(C)$ is equal to

$$F_{\omega_1}(C) := - \int_C^{\omega_2} \frac{du}{-\beta(\omega_1) - \alpha u + \psi(\omega_1, \eta u)}.$$

3 Estimation of the SSEJD

The econometric calibration of the SSEJD model is challenging as financial prices depend upon two unobservable state variables $\delta(t)$ and λ_t . In a first subsection, we describe a sequential Monte-Carlo (SMC) method to filter the state variables when parameters of the SSEJD are known. Its efficiency is checked with simulations. The second subsection develops a first estimation procedure based on a “peaks over threshold” (POT) approach. This method is next applied to fit the SSEJD to a sample of S&P 500 daily data. The third section proposes a particle Monte Carlo Markov Chain (PMCMC) algorithm to improve the quality of the fit. It is based on the SMC filter and is initialized with parameters found by the POT procedure.

3.1 A Particle filter

The Markov chain $\delta(t)$ and the intensity of jumps λ_t are hidden state variables. We use a sequential Monte-Carlo (SMC) method, also called particle filter, to deduce their sample paths. This Bayesian technique is combined later with a Monte-Carlo Markov Chain to estimate the SSEJD, but for the moment, we assume that parameters are known. The estimation approach is based a time-discretization of equations (6), (11) and (5) that define the model. We denote by Δ the length of the time interval. The *ex ante* continuously compounded return (over the period Δ) at time $t_j = j\Delta$, defined by $X_j = \ln S_{(j+1)\Delta} - \ln S_{j\Delta}$, then satisfies the following equation in discrete time

$$X_j = \left(\mu_j - \frac{\sigma_j^2}{2} - \lambda_j \mathbb{E}(e^J - 1) \right) \Delta + \sigma_j \sqrt{\Delta} \varepsilon_j + \Delta L_j \quad (21)$$

where ε_j stands for a standard normal random variable and $\Delta L_j = \sum_{k=N_{j\Delta}}^{N_{(j+1)\Delta}} |J_k|$. Here $N_{(j+1)\Delta} - N_{j\Delta}$ is distributed as a Poisson random variable with parameter $\lambda_j \Delta$. The economic regime is assumed to remain unchanged over the time interval Δ and the value of $\delta(t)$ for $t \in [j\Delta, (j+1)\Delta]$ is denoted by δ_j . The drift and volatility of the Brownian part are constant and respectively equal to $\mu_j = \delta'_j \bar{\mu}$ and $\sigma_j = \delta'_j \bar{\sigma}$.

The Euler approximation of the equation (11) provides the discrete dynamics of the first latent process $\lambda = (\lambda_t)_t$

$$\lambda_{j+1} = \lambda_j + \alpha(\theta_j - \lambda_j)\Delta + \eta \Delta L_j \quad (22)$$

where $\theta_{j-1} = \delta'_{j-1} \bar{\theta}$. It is worth noting that in our applications, Δ is the duration between two consecutive observations and is set to one day. In such a Δ -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 $\Delta L'_j = J\xi_j$ and $\Delta L_j = |J|\xi_j$, where $\xi_j = 1_{\{N_{(j+1)\Delta} - N_{j\Delta} = 1\}}$ is a Bernoulli random variable with probability $\lambda_j \Delta$. The second latent process, δ_j is a discrete Markov chain with the transition matrix $P_\Delta = \exp(Q_0 \Delta)$ and the transition random measure $K(\cdot)$ such that $\delta_{j+1} = \int_{\delta \in E} K(\delta_j, d\delta)$. Remember that at this stage, the model parameters are assumed to be known. We denote by $v_j = (\lambda_j, \delta_j, \Delta L_j)$ the “particle” that puts together all information about the economic regime and the jump process at time $t = j\Delta$. The model admits a useful state-space representation, where the equation (21) provides a measurement equation or system (the ‘space’) that defines the relationship between the (possibly observed) return and the hidden state variables. The vector $v_j = (\lambda_j, \delta_j, \Delta L_j)$ can help finding the transition system (the ‘state’) that describes the dynamics of the state variables. This dynamics depends on the equations (22) and on the dynamics of the Markov chain δ_j .

In the remainder of the paper, we denote by $\{x_1, x_2, \dots, x_n\}$, the sample of observed continuously compounded returns. Conditional to information contained in v_j , the return density $p(x_j | v_j)$ is Gaussian

$p(x_j|v_j) = \mathcal{N}\left(\left(\mu_j - \frac{\sigma_j^2}{2} - \lambda_j \mathbb{E}(e^J - 1)\right) \Delta - \Delta L_j', \sigma_j \sqrt{\Delta}\right)$. On another side, it is possible to simulate the transition density $p(v_{j+1}|v_j)$ with equations (22) and $\delta_{j+1} = \int_{\delta \in E} K(\delta_j, d\delta)$. The density of v_0 is $p(v_0)$ and the posterior distribution of v_j is denoted by $p(v_j | x_{1:j})$. Using standard Bayesian arguments, the expression for the posterior distribution is given by:

$$p(v_j | x_{1:j}) = \frac{p(x_j | v_j)}{\int p(x_j | v_j) p(v_j | x_{1:j-1}) dv_j} p(v_j | x_{1:j-1}) \quad (23)$$

where

$$p(v_j | x_{1:j-1}) = \int p(v_j | v_{j-1}) p(v_{j-1} | x_{1:j-1}) dv_{j-1} \quad (24)$$

To summarize, the calculation of $p(\lambda_j, \delta_j | x_{1:j})$ is done in two steps, close to these used in the Kalman Filter (1960). The first one is a prediction step in which we estimate $p(v_j | x_{1:j-1})$ by the relation (24). In the correction step, we next calculate the probabilities $p(v_j | x_{1:j})$ using the equation (23). In practice, the integral in the prediction step is replaced by a Monte Carlo simulation, of M particles, noted $v_j^{(i)} = (\lambda_j^{(i)}, \delta_j^{(i)}, \Delta L_j^{(i)})$ with importance weights $w_j^{(i)}$, for $i = 1, \dots, M$. The structure of the particle filter algorithm is presented in appendix B. Finally, the filtered intensity for the period j is computed as the sum of particles, weighted by their probabilities of occurrence:

$$\mathbb{E}(\lambda_j | \mathcal{G}_T) = \sum_{i=1:M} \lambda_j^{(i)} w_j^{(i)}$$

and the log-likelihood is approached as follows:

$$\begin{aligned} \log L(\Theta) &= \sum_{j=1}^T \log p(x_j | x_{j-1}) \\ &= \sum_{j=1}^T \log \int p(x_j | v_j) p(v_j | v_{j-1}) dv_j \\ &= \sum_{j=1}^T \log \left(w_j^{(i)} \sum_{i=1}^M p(x_j | v_j^{(i)}) \right) \end{aligned} \quad (25)$$

However, the estimator of the likelihood is not continuous as a function of parameters because it is based on simulations. From a practical viewpoint, maximizing the resulting log-likelihood so as to fit the process is then inefficient. For this reason, we propose an alternative estimation method in the following section.

To illustrate the efficiency of the SMC filter, we firstly simulate a daily sample path of a SSEJD process, with three economic regimes and over a period of five years. The parameters used for this simulation are reported in table 1. The first state corresponds to a period of economic recession: negative average return, high volatility and frequency of jumps. The third regime represents a period of economic growth: positive expected return, low volatility and frequency of jumps. The second state is an intermediate conjuncture, close to economic stagnation. The one year matrix of transition probabilities used in this exercise is presented in table 2.

$\bar{\mu}_1$	-0.05	$\bar{\sigma}_1$	0.30
$\bar{\mu}_2$	0.00	$\bar{\sigma}_2$	0.15
$\bar{\mu}_3$	0.05	$\bar{\sigma}_3$	0.20
$\bar{\theta}_1$	50	ρ^+	50
$\bar{\theta}_2$	25	ρ^-	-60
$\bar{\theta}_3$	10	p	0.5
α	13	η	250

Table 1: This table reports the parameters used for the simulation of a daily sample path of the SSEJD process, with three regimes.

$(p_{ij}(0, 1))_{i,j=1,2,3}$	state 1	state 2	state 3
state 1	0.60	0.20	0.20
state 2	0.20	0.60	0.20
state 3	0.20	0.20	0.60

Table 2: This table presents the one year matrix of transition, used to simulate $\delta(t)$.

M	$\mathbb{E}(\log L(\Theta))$	$std(\log L(\Theta))$	$\frac{std(\log L(\Theta))}{\mathbb{E}(\log L(\Theta))}$
100	34 847	23.05	0.66%
300	35 296	14.68	0.42%
500	35 405	11.65	0.33%
1000	35 513	7.65	0.21%
1500	35 546	5.92	0.17%

Table 3: Average log-likelihoods and standard deviation of log-likelihoods computed with 1000 runs of the particle filter over 1250 simulated observations.

Once that the sample path is simulated, we run the SMC filter with 1500 particles. The graphs of figure 1 compare simulated and filtered intensities of jumps and economic regimes. This tends to confirm the efficiency of the SMC algorithm. We also analyse the averages and standard deviations of log-likelihood estimates for different sizes of the particles set. Next, we run the filter 1000 times and examine the variations of log-likelihood estimates. Results are reported in table 3. We observe that increasing the number of particles reduces significantly the standard deviation of log-likelihoods. However, this noise is small compared to the average estimate, even with 100 particles. This filter is combined in a next section to a Monte Carlo Markov Chain (MCMC) algorithm to estimate the SSEJD. Before, we introduce an approached estimation method that is used to initialize the MCMC algorithm.

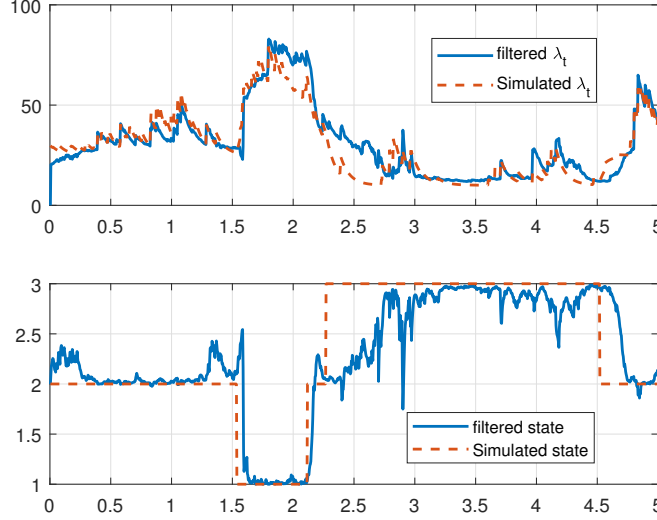


Figure 1: This graph shows simulated and filtered sample paths of λ_t and $\delta(t)$ over 5 years, with parameters of tables 1 and 2.

Notice that there exist more advanced developments of particles filtering like the SMC² of Chopin et al. (2013), the marginalized resample-move approach of Fulp and Li (2013) or the density-tempered marginalized sequential MC samplers. These methods allows to reduce the computational time and can be parallelized.

3.2 A peaks over threshold (POT) procedure

This section proposes an empirical econometric method to estimate SSEJD parameters from time series. This method is used to initialize the Monte Carlo Markov chain algorithm that is detailed in the next section. The procedure counts three stages. In summary, the first step consists to fit a switching diffusion to a sample of data with the Hamilton filter. In the second and third steps, jumps are detected by a peaks over the threshold procedure and parameters defining λ_t are obtained by log-likelihood maximization. All these steps are detailed in the following paragraphs.

The data set is a discrete record of T observations of log-returns, equally spaced with a lag Δ of one day of trading is noted $\{x_1, x_1, x_2, \dots, x_T\}$. In the first stage, we fit to this sample a switching geometric Brownian motion (SGBM) with the filtering technique developed by Hamilton (1989) and inspired from the Kalman filter (1960). The SGBM is used here as an approximation of the SSEJD. In the SGBM, the vector of drift and variance are respectively denoted by $\tilde{\mu} = (\tilde{\mu}_1, \dots, \tilde{\mu}_N)' \in \mathbb{R}^N$ and $\tilde{\sigma} = (\tilde{\sigma}_1, \dots, \tilde{\sigma}_N)' \in \mathbb{R}_+^N$. The N -vector of Gaussian density for x_j in each regime is denoted by $f(x_j)$. The state vector $\delta(t)$ is not directly observable, but the Hamilton filter provides the probabilities of presence in each state. The vector of probabilities of being in a certain regime at time $j\Delta$, $\pi_j = (\pi_j^j)_{j=1 \dots N}$, is computed recursively as follows:

$$\pi_{j+1} = \frac{f(x_j) * (\pi_j' P_\Delta)}{\langle f(x_j) * (\pi_j' P_\Delta), \mathbf{1} \rangle} \quad (26)$$

where $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^N$, $x * y$ is the Hadamard product $(x_1 y_1, \dots, x_N y_N)$ and $P_\Delta = \exp(Q_0 \Delta)$ is the daily matrix of transition probabilities. To start the recursion, we assume that the Markov process $\delta(t)$ has reached its stable distribution. π_0 is then set to the ergodic distribution of $\delta(t)$, which is the

eigen vector of the matrix P_Δ , coupled to the eigen value equal to 1. If we dispose of T observations, the log-likelihood function is:

$$\ln l(x_1, x_1, x_2, \dots, x_T) = \sum_{j=1}^T \ln \langle f(x_j), (\pi_{j-1} P_\Delta) \rangle. \quad (27)$$

This log-likelihood is maximized in order to estimate the SGBM parameters. In a second stage, we use the peaks over threshold method to detect jumps. We postulate that whether the observed return is above or below some thresholds, it is likely that a jump occurred. The thresholds are noted $g(\alpha_1)$ and $g(\alpha_2)$ where α_1, α_2 are the chosen confidence levels. If $\Phi(\cdot)$ denotes the pdf of a standard normal, $g(\alpha_1), g(\alpha_2)$ are set respectively to the α_1 and α_2 percentiles of the normal distribution fitted to the sample:

$$g(\alpha_k) = \sqrt{\Delta} \hat{\sigma} \Phi^{-1}(\alpha_k) \quad k = 1, 2$$

where $\hat{\sigma}$ is the standard deviation of the whole sample. The Gaussian distribution being not heavy-tailed, detecting large jumps requires to choose percentiles located far in the tails, as suggested in Embrechts et al. (2011). For this reason, we choose the 5% and 95% percentiles as threshold.³ There probably exist alternative ways to improve the detection of jumps with a POT approach in a switching regime models. However as our purpose is mainly to find a relevant initial set of parameters for the MCMC algorithm, we haven't explored them.

Once that jumps are detected, the dynamics of the stocks index is approached as follows:

$$\begin{aligned} (x_j - \tilde{\mu}_j^{ML} \Delta) &\sim J_j \text{ if } (x_j - \tilde{\mu}_j^{ML} \Delta) > g(\alpha_1) \text{ or } (x_j - \tilde{\mu}_j^{ML} \Delta) < g(\alpha_2) \\ x_j &\sim \tilde{\mu}_j^{ML} \Delta + \tilde{\sigma}_j^{ML} W_\Delta \quad \text{otherwise} \end{aligned}$$

where μ_j^{ML} and $\tilde{\sigma}_j^{ML}$ are respectively the drift and standard deviations of the most likely state (detected with the Hamilton filter) at time j . Jump parameters (ρ^-, ρ^+, p) are then fitted by log-likelihood maximization of the jump probability density function:

$$(\rho^-, \rho^+, p) = \arg \max \sum_{i=1}^T \log \nu(x_i | \rho^-, \rho^+, p) I_{jump \text{ at } t_i}.$$

At this stage, it remains to estimate the parameters involved in the dynamics of λ_t . For this purpose, we construct the discretized sample path of λ_t . If we denote the variations of intensity over a time interval by $\Delta\lambda_i = \lambda_i - \lambda_{i-1}$, for given estimates of $\theta_{1...N}$, α and η , we have that

$$\Delta\lambda_j = \alpha(\theta_j^{ML} - \lambda_{j-1})\Delta + \eta J_j I_{jump \text{ at } t_j}.$$

where θ_j^{ML} is the mean reversion level corresponding to the most likely state. $I_{jump \text{ at } t_j}$ is an indicator variable equal to one if a jump is detected at time t_j . If we set λ_0 to its most likely mean reversion level at time zero, the values of intensities at times $i\Delta$ are obtained by summing up the variations:

$$\lambda_i = \lambda_0 + \sum_{j=1}^i \Delta\lambda_j \quad i = 0, \dots, n$$

³We can think to relate thresholds to regimes. For example, if we denote by $\tilde{\sigma}_i^{ML}$ the volatility of the SGBM in the most likely state at time t_i , we can assume that $g(\alpha_k, i) = \sqrt{\Delta} \tilde{\sigma}_i^{ML} \Phi^{-1}(\alpha_k) \quad k = 1, 2$. However, when this method is applied to the S&P 500 data set, fewer jumps are detected during recessions than in periods of growth. This counter intuitive result is explained by the fact that thresholds are proportional to the volatility in each regime. As the volatility is much important during recessions than in other periods, thresholds are also much higher. Consequence: less log-returns exceed thresholds during bad economic times. This observation motivates us to not relate thresholds to regimes and to use instead the standard deviation of the whole sample. This reduces the accuracy of the POT method. However, given that we use it to find an acceptable starting point for the MCMC algorithm, this loss of accuracy has a limited impact on final conclusions.

When Δ is small, the probability of observing a jump in the i^{th} interval of time is equal to $\lambda_i \Delta$. The set of parameters defining the dynamics of intensity is obtained by maximization of log-likelihood:

$$(\alpha, \eta, \theta_{1...N}) = \arg \max \sum_{i=1}^n \log ((\lambda_i \Delta) I_{jump \text{ at } t_i} + (1 - \lambda_i \Delta) I_{no \text{ jump at } t_i}) \quad (28)$$

Finally, we compute the vector of volatility $\bar{\sigma}$, with the sample from which jumps are removed and under the assumption that the most likely regime is the observed one.

To illustrate this procedure, We collect a sample of S&P 500 daily data, from May 2000 to May 2016. The time series contains a total of 4023 returns. Table 1 provides associated statistics. The yearly volatility measured over this period is 19.76% and the very high sample kurtosis indicates that the distribution of returns has fat-tails. Jarque Bera and Lillie tests reject normality whereas the Durbin Watson statistic reveals serial dependence.

	Value
Mean daily return	0.01%
Standard daily deviation	1.25%
Skewness	-0.18
Kurtosis	11.28
Jarque Bera p-value	1e-3
Lillie test p-value	1e-3
Durbin Watson p-value	0.000

Table 4: 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 May 2000 to May 2016.

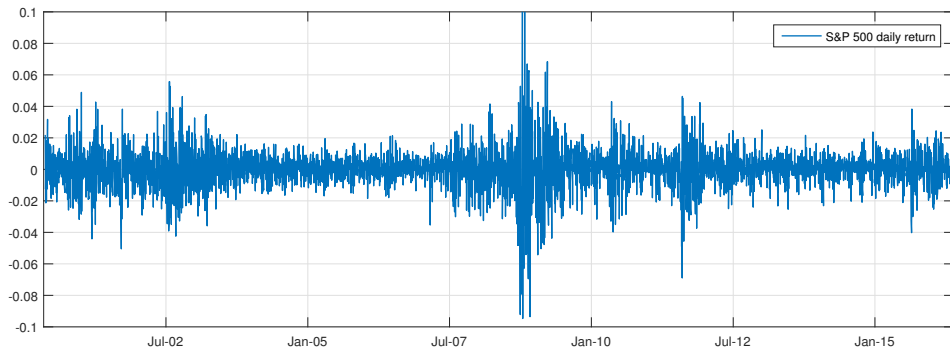


Figure 2: S&P 500 daily return from May 2000 to May 2016.

Figure 2 plots returns of the S&P index over the sample period. From May 2000 to end 2004, the S&P is volatile and we observe several trains of grouped jumps, related to the internet bubble burst, to the terrorist attack of the 11th of September 2001 and to the US invasion of Iraq in 2003. From end 2004 to July 2007, the volatility falls and we don't observe any shock of big amplitude. Clustering of large jumps

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 during these periods do not display any clear trend: negative movements alternate regularly with large positive moves. Finally, from mid 2015 to May 2016, the uncertainty about the recovery of the US economy and the fear of deflation in the Eurozone raise the volatility.

The table 4 reports the parameters of the switching geometric Brownian motion, with three regimes, fitted to the time series of the S&P 500. The first state corresponds to periods of strong economic growth. The volatility of stocks is small and stocks earn on average 23.5%. The third regime is assimilated to a severe economic recession: the volatility is above 40% and the average stocks return is -38%. The second regime is an intermediate economic state of stagnation.

$\tilde{\sigma}_1$	9.24%	$\tilde{\mu}_1$	23.46%
$\tilde{\sigma}_2$	18.21%	$\tilde{\mu}_2$	-8.10%
$\tilde{\sigma}_3$	42.28%	$\tilde{\mu}_3$	-38.72%
$(p_{ij}(0, 1 \text{ day}))_{i,j=1,2,3}$	state 1	state 2	state 3
state 1	0.9775	0.0225	0.0000
state 2	0.0225	0.9714	0.0061
state 3	0.0000	0.0246	0.9754
Log.Lik.	12 615	AIC	-25 191
BIC	-25 065		

Table 5: This table reports the parameters of a switching geometric Brownian motion, fitted with the Hamilton filter to the S&P 500 time series.

In the introduction, we mention that self-excited jump processes fail to replicate economic cycles. Mainly because self-excited shocks occur during economic recessions and disappear in periods of economic growth. To emphasize this point, we fit by the POT procedure a one dimension self-excited jump diffusion (SEJD) and a 3 dimensions SSJED. The chosen confidence levels are $\alpha_1 = 5\%$ and $\alpha_2 = 95\%$. Tables 6 and 7 report the estimated parameters and the log-likelihoods, computed with the particle filter.

An analysis of AIC and BIC confirms that the SSEJD fits the data set better than the SGBM and the SEJD. A comparison of $\tilde{\sigma}$ and $\bar{\sigma}$ allows us to evaluate the part of the volatility caused by the jump process. In the first regime, as $\tilde{\sigma}_1$ and $\bar{\sigma}_1$ are nearly identical, the jump process is quasi inactive. In the second and third regimes, jumps account for respectively 4% and 25% of the total volatility. This emphasizes that self-excited jumps mainly occur during period of moderate or severe economic turmoil. Ignoring this aspect leads to underestimate the level toward the jumps arrival intensity revert. Indeed, the estimate for θ in the SEJD is only 3.49 whereas it climbs to 38.47 in the third regime of the SSJED.

$\bar{\sigma}_1$	12.78%	$\bar{\mu}_1$	2.32%
θ	3.49	ρ^+	31.24
α	17.48	ρ^-	-31.99
η	553.36	p	0.4848
Log.Lik.	12 559	AIC	-25 103
BIC	-25 052		

Table 6: This table reports the parameters of the 1D SEJD calibrated with the POT method ($\alpha_1 = 5\%$ and $\alpha_2 = 95\%$) applied to the S&P 500 time series.

$\bar{\sigma}_1$	10.05%	$\bar{\mu}_1$	23.46%
$\bar{\sigma}_2$	14.20%	$\bar{\mu}_2$	-8.10%
$\bar{\sigma}_3$	17.50%	$\bar{\mu}_3$	-38.72%
$\bar{\theta}_1$	2.29	ρ^+	31.24
$\bar{\theta}_2$	9.60	ρ^-	-31.99
$\bar{\theta}_3$	38.47	p	0.4848
α	30.06	η	645.09
Log.Lik.	12 745	AIC	-25 449
BIC	-25 323		

Table 7: This table reports the parameters of the 3D SSEJD calibrated with the POT method ($\alpha_1 = 5\%$ and $\alpha_2 = 95\%$) applied to the S&P 500 time series.

3.3 Calibration by Particle Monte Carlo Markov Chain

When dealing with a non-Gaussian and nonlinear specification, simulation-based methods offer strong advantages over the alternative approaches. Among the most popular 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)). To estimate a jump-diffusion process from equity returns, Andersen, Benzoni, and Lund (2002) use an EMM approach. Eraker et al. (2003) develop a likelihood-based estimation approach for estimating multivariate jump-diffusion using Markov Chain Monte Carlo (MCMC) methods.

In this paper we employ a Particle Markov Chain Monte Carlo method (PMCMC) to fine tune parameters obtained the POT procedure. We refer to Doucet et al. (2000) for a review of other simulation-based methods. The set of parameters is denoted by Θ in the sequel and serves us as index for the probability distribution function. We adopt a Bayesian approach to estimate Θ by computing the parameters posterior distribution

$$\pi(\Theta) = p(\Theta | x_{1:T}) = \frac{p(\Theta)p(x_{1:T} | \Theta)}{\int p(\Theta')p(x_{1:T} | \Theta')d\Theta'} \quad (29)$$

where $p(\Theta)$ and $p(x_{1:T} | \Theta)$ denotes respectively the parameter prior distribution and the likelihood of the data. The density $\pi_\Theta(\Theta)$ is built by the PMCMC method that generates a sample from $\pi_\Theta(\Theta)$ by creating a Markov chain with the same stationary distribution as the parameters posterior one. Once that the Markov chain has reached stationarity after a transient phase, called “burn-in” period, samples from the

posterior distribution can then be simulated. Standard MCMC algorithm requires a point-wise estimate of $p(x_{1:T}|\Theta)$, that is not available in our model. Instead, $p(x_{1:T}|\Theta)$ is approached by its estimate yield by a particle filter. The construction of the Markov chain consists of two steps, repeated iteratively.

At the beginning of the k^{th} iteration, we propose a candidate parameter Θ' from a proposal distribution $q(\Theta'|\Theta^{(k-1)})$ given the previous state of the Markov chain, noted $\Theta^{(k-1)}$. The proposal distribution has a support that covers the target distribution. In the second step, we determine if we update the state

by Θ' . For this purpose, the acceptance probability of the Metropolis-Hastings algorithm is computed as follows

$$\varepsilon(\Theta', \Theta^{(k-1)}) = \min \left\{ 1, \frac{\pi(\Theta')}{\pi(\Theta^{(k-1)})} \frac{q(\Theta^{(k-1)} | \Theta')}{q(\Theta' | \Theta^{(k-1)})} \right\}. \quad (30)$$

This determines the probability that we assign the candidate parameter as the next state of the Markov chain, $\Theta' \rightarrow \Theta^{(k)}$. Intuitively, if we disregard the influence of the proposal q , a candidate is accepted if it increases the posterior likelihood $\pi(\Theta') > \pi(\Theta^{(k-1)})$. The presence of $q(\cdot)$ in equation (30) allows a small decrease in the posterior likelihood, so as to explore the entire posterior.

The resulting K samples $\Theta^{(1:K)}$ (after the burn in period) serve next to build the empirical distribution of $\pi(\Theta)$, which is defined by

$$\hat{\pi}(\Theta) = \frac{1}{K} \sum_{k=1}^K \delta_{\Theta^{(k)}}(d\Theta),$$

where $\delta_{\Theta^{(k)}}(d\Theta)$ are the Dirac atoms located at $\Theta = \Theta^{(k)}$, with equal weights. The expected parameters with respect to the posterior distribution of parameters is then approached as follows

$$\mathbb{E}(\Theta|x_{1:T}) \approx \frac{1}{K} \sum_{k=1:K} \hat{\pi}(\Theta^{(k)}) \Theta^{(k)}$$

In numerical applications, the transition distribution $q(\Theta' | \Theta^{(k-1)})$ is assumed Normal, $\mathcal{N}(\Theta' | \Theta^{(k-1)}, \sigma_q)$ of mean $\Theta^{(k-1)}$ and standard deviation σ_q . As this distribution is symmetric, $q(\Theta^{(k-1)} | \Theta') = q(\Theta' | \Theta^{(k-1)})$, the acceptance probability simplifies to

$$\begin{aligned} \varepsilon(\Theta', \Theta^{(k-1)}) &= \min \left\{ 1, \frac{\pi(\Theta')}{\pi(\Theta^{(k-1)})} \right\} \\ &= \min \left\{ 1, \frac{p(\Theta')p(x_{1:T} | \Theta')}{p(\Theta^{(k-1)})p(x_{1:T} | \Theta^{(k-1)})} \right\}. \end{aligned}$$

We use the PMCMC algorithm to improve the calibration of the SEJD and the 3D-SSEJD model to the S&P 500 time serie. This algorithm is applied to the set of parameters $\Theta = \{\bar{\sigma}, \bar{\theta}, \alpha, \eta, p, \rho^+, \rho^-, \bar{\mu}, (q_{i,j})_{i \neq j}\}$. If I_n is a n -vector of ones, $q(\Theta' | \Theta^{(k-1)})$ is assumed Normal, $\mathcal{N}(\Theta' | \Theta^{(k-1)}, \sigma_q)$ with a standard deviation σ_q set to:

$$\sigma_q = (0.5\% \times I_3, 0.1\% \times I_3, 1, 1, 0.04, 1, 1, 0.5\% \times I_3, 0.1\% \times I_6)$$

Initial values for Θ are these obtained with the POT method (tables 6 and 7). The prior distribution is a normal distribution whose standard deviation equal to 15% of initial parameters. The filter runs with 500 particles and we perform 10 000 iterations of the PMCMC procedure.

We obtain an acceptance rate around 30%⁴. The convergence is confirmed by graphs of figure 3 that present log-likelihoods computed by the particle filter during the PMCMC loop, for the 3D SSEJD. Log-likelihood estimates becomes stable after a burn in period of 5000 iterations. Notice that oscillations are caused by the recalculation of the log-likelihood estimate at each step of the algorithm, even in case of rejection of the candidate set of parameters. We also tried the PMCMC algorithm with arbitrary but realistic initial values for Θ , the loglikelihood still converges but the speed of convergence is much lower than with prior POT estimates.

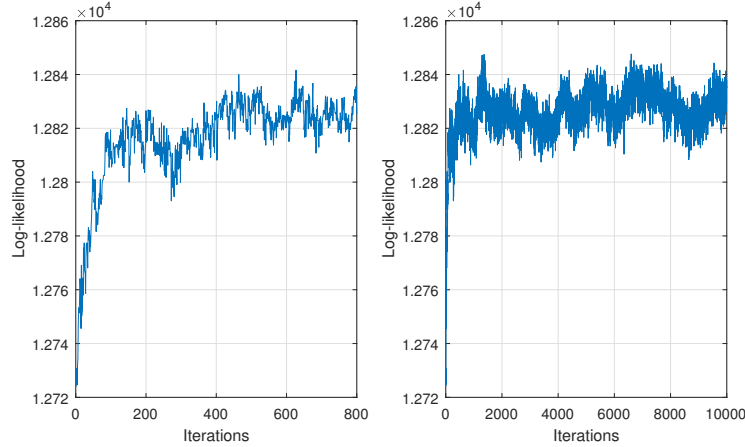


Figure 3: Evolution of log-likelihood computed by the particle filter during the fit of a 3D SSEJD to the S&P 500 time series. The left plot presents the first 800 iterations. The right plot shows the 10 000 iterations.

Parameters estimated by this way are reported in table 8, 9 and 10. A comparison of log-likelihoods, BIC and AIC with these obtained with the POT method, confirms that the PMCMC algorithm improves the quality of the fit for the 1D SEJD and 3D SSEJD. As previously the first, second and third regimes respectively correspond to periods of economic growth, of moderate recession and of serious economic turmoil.

	Estimate	St.dev.		Estimate	St.dev.
$\bar{\sigma}_1$	11.72%	0.60%	$\bar{\mu}_1$	11.18%	4.16%
$\bar{\theta}_1$	6.89	1.88	η	549.24	11.28
ρ^+	75.85	9.48	α	8.52	0.85
ρ^-	-69.18	8.39	p	0.38	0.12
Log.Lik.	12 667		AIC	-25 318	
BIC	-25 267				

Table 8: This table reports the averages and standard deviations of 1D SEJD parameters fitted with the PMCMC procedure.

⁴In theory the acceptance rate can be improved to any desired level. Increasing the number of particles raises considerably the computational time of the estimation procedure.

	Estimate	St.dev.		Estimate	St.dev.
$\bar{\sigma}_1$	9.24%	0.42%	$\bar{\mu}_1$	12.35%	5.21%
$\bar{\sigma}_2$	17.73%	0.94%	$\bar{\mu}_2$	-7.84%	4.98%
$\bar{\sigma}_3$	36.38%	3.66%	$\bar{\mu}_3$	-43.15%	11.72%
$\bar{\theta}_1$	2.92	1.38	η	675.95	9.67
$\bar{\theta}_2$	3.60	1.36	α	29.37	9.39
$\bar{\theta}_3$	39.40	1.90	p	0.36	0.21
ρ^+	72.38	9.05	ρ^-	-58.34	13.32
Log.Lik.	12 837		AIC	-25 635	
BIC	-25 509				

Table 9: This table reports the averages and standard deviations of 3D SSEJD parameters fitted with the PMCMC procedure.

$(p_{ij}(0, 1 \text{ day}))_{i,j=1,2,3}$	state 1	state 2	state 3
state 1	0.9752	0.023	0.0019
state 2	0.0234	0.9706	0.0060
state 3	0.0017	0.0242	0.9741
Matrix Q	$q_{.,1}$	$q_{.,2}$	$q_{.,3}$
$q_{1,.}$	-6.3496	5.8929	0.4568
$q_{2,.}$	6.0030	-7.5432	1.5402
$q_{3,.}$	0.3662	6.2149	-6.5810
St. Dev. Q	$q_{.,1}$	$q_{.,2}$	$q_{.,3}$
$q_{1,.}$	0.1505	0.0482	0.1407
$q_{2,.}$	0.1000	0.1328	0.0511
$q_{3,.}$	0.0416	0.0605	0.0648

Table 10: This table reports the matrix of daily transition probabilities, the matrix of instantaneous probabilities Q and standard deviations of estimates.

Figures (4) and (5) display the filtered jump intensities for the SEJD and the 3D SSEJD, with parameters estimated by the PMCMC algorithm. In the SEJD model, the jump intensity is nearly null from 2003 to mid 2007 and from mid 2012 to 2015, that correspond to periods of economic growth in the US. This confirms that self-excited shocks are foremost observed during periods of economic turmoil. The long periods of low jumps activity drive down the reversion level of the SEJD to $\theta = 6.89$. This level is particularly low compared to the value reached by the intensity during the credit crunch crisis. In this sense, the realization of the filtered sample path of λ_t seems very unlikely.

The analysis of filtered states variables of the 3D SSEJD reveals that the Markov chain is in state 3 during the Iraqi invasion of 2002 and the credit crunch of 2008. The intensity of jumps reaches its highest values (above 60) in September 2008, precisely after the fall of Lehman Brothers. The realization of such a maximum is more likely than in the SJED, as the intensity increases on average by $\eta \left(\frac{p}{\rho^+} - \frac{1-p}{\rho^-} \right) = 10.78$ at each jumps and as λ_t reverts to a high frequency ($\theta_3 = 39.40$) in regime 3. Between end 2004 to July 2007 and from mid 2012 to 2015, $\delta(t)$ is in the first regime. The filtered intensity in this regime is nearly null and close to the mean reversion level $\bar{\theta}_1 = 2.92$. The rest of the time, the chain is in the second

state and the jumps activity is moderate. These observations confirm our previous intuition: ignoring the influence of economic regimes on the dynamics of jumps lead to underestimate the reversion level of λ_t during economic recessions.

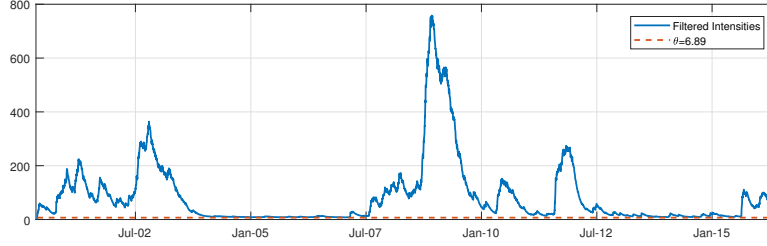


Figure 4: Graph of filtered intensities for the 1D SEJD, with the particle filter ($M = 3000$) and parameters of table 8.

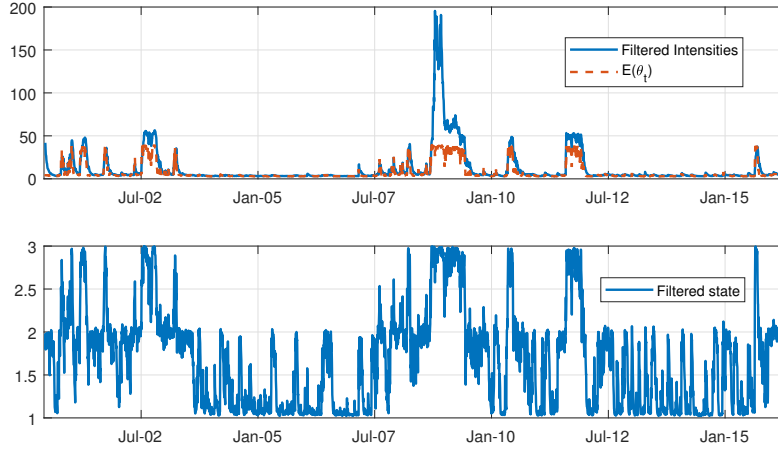


Figure 5: The upper graph plots filtered intensities of the 3D SSEJD. The lower graph shows the average state of the Markov Chain $\delta(t)$. Both are computed with the particle filter ($M = 3000$) and parameters of table 9.

To conclude this section, we study the first moments of the intensity with results of proposition 2.1, 2.2 and 2.3. The left graph of figure 6 shows the expected intensity computed with parameter estimates of table 9 and $\lambda_0 = 10$, for different initial states of $\delta(0)$. We clearly observe the convergence of intensities whatever the initial economic regime to $\lambda_\infty = (12.8, 12.8, 12.8)^\top$, the asymptotic limit of $\mathbb{E}(\lambda_t|\mathcal{F}_0)$ obtained with the equation (14). During the first two months, expected intensities quickly moves during the first year toward their mean reversion level $\theta_{i=1,2,3}$. We explain this by the high speed of mean reversion ($\alpha = 29.37$) dragging λ_t to $\theta_{i=1,2,3}$. Next $\mathbb{E}(\lambda_t|\mathcal{F}_0)$ converges to λ_∞ . The right plot of figure 6 shows the standard deviation of λ_t . Volatilities as expectations, converge toward the same value (around 21.41), whatever the initial state. In regime 3, the curve of standard deviation is humped decreasing whereas in regime 1 and 2, this curve is concave and strictly increasing.

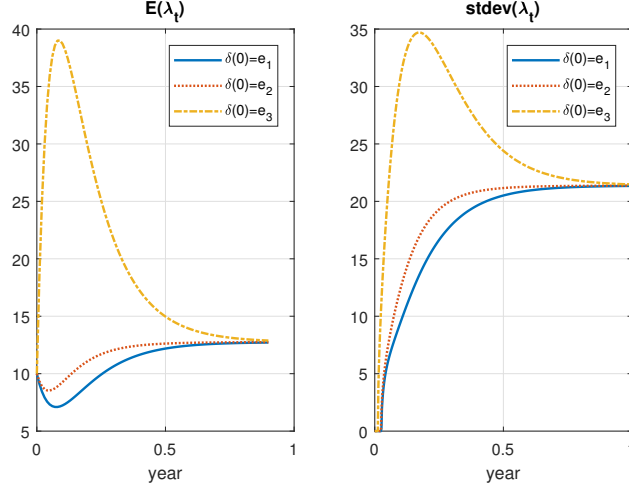


Figure 6: These graphs shows expectations and standard deviations of λ_t , for different initial regimes. λ_0 is set to 10. Parameters are those reported in table 9.

	3 D model			1 D model
	$\delta(0) = e_1$	$\delta(0) = e_2$	$\delta(0) = e_3$	
$P(N_1 = 0)$	1.09%	0.96%	0.15%	0.02%
$P(N_1 = 1)$	3.15%	2.81%	0.54%	0.10%
$P(N_1 = 2)$	10.42%	9.44%	2.20%	0.54%
$P(N_1 = 3)$	37.27%	34.09%	10.01%	3.08%
$P(N_1 \geq 4)$	48.06%	52.71%	87.10%	96.26%

Table 11: Comparison of distributions of N_t after 1 year for the non-switching (1D) and switching models (3D). $\lambda_0 = 20$. Parameters are those reported in table 9.

Table 11 reports the probabilities of observing zero to three jumps and above over one year for the one and three regime models. These are computed by numerical derivation of the probability generating function of N_t of proposition 2.4. In the 3D model when $\delta(t) = e_1$, the intensity quickly reverts to a low mean reversion level ($\theta_1 = 2.92$) and the probability of observing three jumps or less climbs to 51.94%. Given that the mean reversion level in the second regime is around 3.60, the probability of observing three jumps or less falls to 47.29%. Whereas, this probability falls to 12.9% if $\delta(0)$ is in the third state. This confirms that self-excited shocks occur mainly during economic recessions (state 3) and are rare events in other economic regimes. A comparison with the non-switching model emphasizes that ignoring economic regimes leads to underestimate the probability of a shock during financial turmoil and to overestimate the jump risk during economic growth.

4 Validation

In this section, we estimate the 3D SSEJD on simulated data in order to validate previous results. We simulate a sample path of 3750 daily log-returns for a 3D SSEJD. Parameters used in this simulation are reported in the first column of table 12. The log-likelihood of this sample path, computed by the SMC procedure with 3000 particles is equal to 10 974. The second column of table 12 reports the drifts, standard deviations and transition probabilities of a 3D switching Brownian motion fitted to our sample path with the Hamilton filter. Estimates of volatilities and transition probabilities are close to parameters

used for the simulation. Next, we apply the POT method in order to fit parameters of the switching jump component. Results of this procedure are reported in the second column of table 12. Estimates of η , α and p are not far from real values. However, the amplitude of jumps and volatilities are respectively over and under estimated. The reason is that we ignore the diffusion part for days during which a jump is detected by the POT method. This last set of parameters serve us as starting point for the PMCMC algorithm. As in the previous section, we run 10 000 steps of this procedure with 500 particles. The right graph of figure 9 emphasizes that the log-likelihood quickly converges toward a level that is close to the log-likelihood evaluated with real parameters.

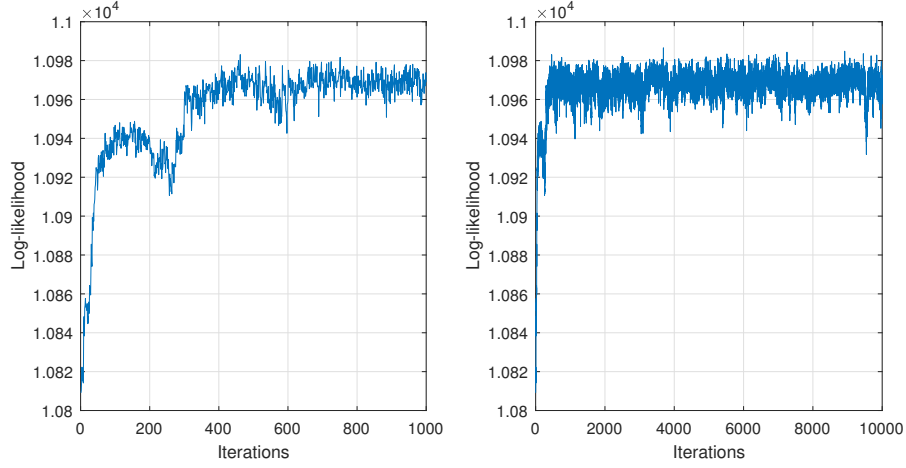


Figure 7: Evolution of log-likelihood computed by the particle filter during the fit of a 3D SSEJD to a simulated sample path. The left plot shows the first 1000 iterations whereas the right plot presents all iterations.

The last two columns of table 12 reports the average and standard deviation of parameters, computed over the last 5000 iterations. Of course, we do not retrieve exactly the parameters used to generate the sample path. However, estimates are in the range of these initial parameters. In particular, estimates of $(\sigma_i)_{i=1,2,3}$, $(\theta_i)_{i=1,2,3}$, η , α , p and $(q_{ij})_{ij=1,2,3}$ are close to their original values. The estimates of μ_1 , ρ^+ and ρ^- are less accurate. Finally, we run the SMC procedure with 3000 particles and using PMCMC parameter estimates. The figure 8 compares filtered and simulated sample paths for λ_t and $\delta(t)$. Their similarity is striking and validates our calibration approach.

Parameters	Initial	SBM	POT	PMCMC	std. Dev.
$\bar{\sigma}_1$	10.00%	9.85%	10.68%	9.77%	0.47%
$\bar{\sigma}_2$	20.00%	21.14%	16.97%	20.70%	0.91%
$\bar{\sigma}_3$	30.00%	33.59%	19.87%	32.06%	1.78%
$\bar{\theta}_1$	5	-	2.56	4.54	1.20
$\bar{\theta}_2$	15	-	19.30	12.61	1.99
$\bar{\theta}_3$	30	-	42.49	30.93	3.08
$\bar{\mu}_1$	10.00%	20.21%	20.21%	16.61%	9.44%
$\bar{\mu}_2$	0.00%	3.29%	3.29%	0.03%	5.28%
$\bar{\mu}_3$	-10.00%	-10.63%	-10.63%	-10.98%	5.48%
η	500	-	434.36	472.14	18.80
α	50	-	60.35	46.13	10.13
p	0.5	-	0.4934	0.4551	0.2520
ρ^-	-80	-	-30.14	-95.84	12.03
ρ^+	80	-	30.37	89.49	10.18
q_{12}	5	4.4753	4.4753	4.4963	0.0506
q_{13}	0	0.000	0.000	0.0028	0.0248
q_{21}	3	2.3080	2.3080	3.3868	0.0602
q_{23}	2	2.8976	2.8976	2.5689	0.1464
q_{31}	0	0.0982	0.0982	0.0458	0.1370
q_{32}	6	4.9472	4.9472	4.7940	0.0350
Log. Like.	10 974	10 775	10 820	10 968	

Table 12: This table reports 3D SSEJD parameters and their estimates obtained with the POT and PMCMC methods.

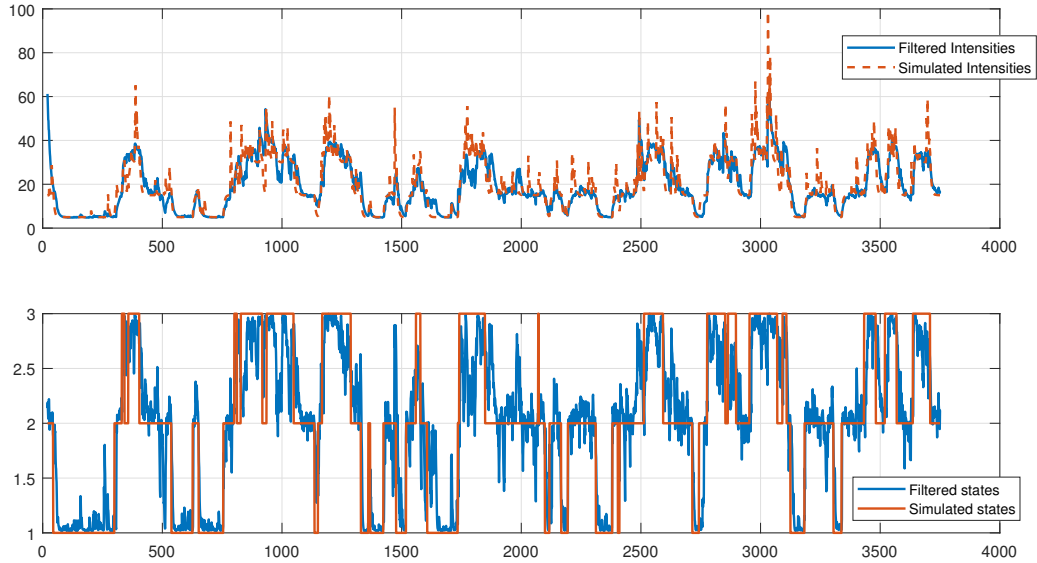


Figure 8: This graph shows simulated and filtered sample paths of λ_t and $\delta(t)$ over 3750 days. Simulated data are obtained with parameters in the first column of table 12.

5 Change of measures and derivatives pricing

The previous sections emphasize that the SSEJD model is appropriate for econometric modelling. We show in this section that it is also adapted for pricing purposes. A fundamental step in derivatives pricing is the definition of a risk-neutral measure. To avoid arbitrage opportunities, the price of any contingent claim asset is evaluated by the expected present value of the future net cash flow under an equivalent measure \mathbb{Q} . Under this measure, all assets earn on average the risk free rate. However, in our framework the financial market is incomplete by nature and there is no unique risk neutral measure. Different equivalent measures exist like the minimal entropy martingale measure or the variance-optimal martingale measure. Although the selection of equivalent measures itself is an important research topic in finance, this paper will not discuss which criterion is the best. Instead, we use the Esscher transformation method. The Esscher transform has been promoted by Gerber and Shiu (1994) in option pricing. The merit of the risk neutral measure determined by the Esscher transform is that it provides a general, transparent and unambiguous framework. As shown in Elliott et al. (2005), a regime-switching Esscher transform is defined by a so-called regime-switching Esscher parameter $\xi_t = \delta(t)' \xi$, where $\xi = (\xi_1, \xi_2, \dots, \xi_N)'$. Specifically, the risk neutral measure \mathbb{Q} is equivalent to \mathbb{P} through the Radon Nykodym density as follows

$$\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_t} = \frac{e^{\int_0^t \xi_s dW_s}}{\mathbb{E}^P \left[e^{\int_0^t \xi_s dW_s} \mid \mathcal{G}_t \right]} = e^{-\frac{1}{2} \int_0^t \xi_s^2 \sigma_s^2 ds + \int_0^t \xi_s \sigma_s dW_s}.$$

Since ξ and σ are bounded, the stochastic exponential on the right hand side of the above equation is an $(\mathcal{F}, \mathbb{P})$ -martingale and then the equivalent measure \mathbb{Q} is well defined. The risk free rate is assumed constant and equal to r . The next proposition gives the dynamics of $\{S_t\}_{t \geq 0}$ under \mathbb{Q} :

Proposition 5.1. *If the vector of parameters ξ is defined by*

$$\xi_j = \frac{\bar{\mu}_j - r}{\bar{\sigma}_j^2} \quad j = 1, \dots, N$$

then the process $\{S_t\}_{t \geq 0}$ is governed by the following SDE under \mathbb{Q}

$$\frac{dS_t}{S_{t-}} = rdt + \sigma_t dW_t^Q + d \left(\sum_{j=1}^{N_t} (e^{J_j} - 1) \right) - \lambda_t \mathbb{E} (e^J - 1) dt, \quad (31)$$

where

$$W_t^Q = W_t - \int_0^t \xi_s \sigma_s ds, \quad \forall t \geq 0, \quad (32)$$

is a Brownian motion under \mathbb{Q}

Proof. From Girsanov's theorem, the process defined by (32) is a standard Brownian motion under \mathbb{Q} . Substituting (32) into (6) immediately gives the dynamics of X_t under \mathbb{Q} and its drift, i.e. (31). \square

In the remainder of this section, we summarize the method to price an European derivative of maturity T , written on S_t . A call option is considered with a payoff and a strike expressed as a function of the log-return $\ln(\frac{S_T}{S_0})$, and of k , such that the strike is equal to $K = S_0 e^k$. The price of the call option is then a function of the log- strike $C(k)$. The values of states variables are here considered as extra parameters.

If the risk neutral density at time $t \leq T$ of the log return $\ln \frac{S_T}{S_0} | \mathcal{F}_t$ for a given couple of values $(\lambda_0, \delta(0))$ is denoted by $f_{t,T}(x)$, option prices are equal to their expected discounted payoffs

$$C(k) = S_0 \int_k^{+\infty} e^{-r(T-t)} (e^x - e^k) f_{t,T}(x) dx,$$

As $C(k)$ tends to S_t (resp. $-S_t$) whether $k \rightarrow -\infty$, $C(k)$ is not square integrable with respect to k and their Fourier transforms are not defined. For this reason, we consider the modified call price denoted by $c(k) = e^{\epsilon k} C(k)$ and for which the Fourier transform exists for some ϵ ($\epsilon > 1$ for the call). The Fourier transforms of $c(k)$ is defined as follows:

$$\mathcal{FC}(\omega) = \int_{-\infty}^{\infty} e^{i\omega k} c(k) dk,$$

A direct calculation leads to:

$$\mathcal{FC}(\omega) = \frac{S_0 e^{-r(T-t)}}{(i\omega + \epsilon)^2 + (i\omega + \epsilon)} \Upsilon_{t,T}(i\omega + \epsilon + 1),$$

If the mgf of the log return $\mathbb{E}(e^{\omega X_T} | \mathcal{F}_t)$, as introduced in the proposition 2.5, is denoted $\Upsilon_{t,T}(\omega)$, the values of call options are obtained by inverting the Fourier transform:

$$C(k) = \frac{S_0 e^{-\epsilon k - r(T-t)}}{\pi} \int_0^{\infty} e^{-i\omega k} \frac{\Upsilon_{t,T}(i\omega + \epsilon + 1)}{(i\omega + \epsilon)^2 + (i\omega + \epsilon)} d\omega, \quad (33)$$

The naive approach consists to calculate numerically the integral present in equations (33). Setting $\omega_j = \Delta_\omega(j-1)$, an approximation of the call price is in this case given by:

$$C(k) \approx \frac{S_0 e^{-\epsilon k - r(T-t)}}{\pi} \sum_{m=1}^M e^{-i\omega_m k} \delta_m \left[\frac{\Upsilon_{t,T}(i\omega_m + \epsilon + 1)}{(i\omega_m + \epsilon)^2 + (i\omega_m + \epsilon)} \right] \Delta_\omega \quad (34)$$

where $\delta_j = \frac{1}{2}1_{\{j=1\}} + 1_{\{j \neq 1\}}$. An adapted choice for the discretization steps in the equation (34), allows us to use a Fast Fourier Transform algorithm to speed up calculations. This point is detailed in the next proposition

Proposition 5.2. *Let M be the number of steps used in the Discrete Fourier Transform (DFT) and $\Delta_k = \frac{2k_{max}}{M-1}$, be the step of discretization. Let us denote $\delta_j = \frac{1}{2}1_{\{j=1\}} + 1_{\{j \neq 1\}}$, $\Delta_\omega = \frac{2\pi}{M\Delta_k}$ and $\omega_j = (j-1)\Delta_\omega$. The values of $C(k)$ at points $k_j = -\frac{M}{2}\Delta_k + (j-1)\Delta_k$ are approached by*

$$C(k_j) \approx \frac{2S_0 e^{-\epsilon k - r(T-t)}}{M\Delta_k} \text{Re} \left(\sum_{m=1}^M \delta_m \left(\frac{\Upsilon_{t,T}(i\omega_m + \epsilon + 1)}{(i\omega_m + \epsilon)^2 + (i\omega_m + \epsilon)} (-1)^{m-1} \right) e^{-i\frac{2\pi}{M}(m-1)(j-1)} \right) \quad (35)$$

This last relation can be computed by any fast Fourier transform algorithm.

To illustrate this section, we price European calls on the S&P 500. The maturities range from 1 to 12 months and strike prices run from 95% to 105% of the spot S&P value. We use the parameters reported in table 9. The first graph of figure 9 shows the surface of implied volatilities in the second state. For short term maturities, we observe a strong asymmetry of the smile of volatility. For longer maturities, the smile is flatter between 17% to 33%, depending upon the economic regime. The second graph plots the 6 months smile for $\delta(0) = e_1$ to e_3 with $\lambda_0 = \bar{\theta}'\delta(0)$. In state 3, that correspond to a depressed economic conjuncture, the smile is flatter but the average level of implied volatilities is much higher than in regimes 1 and 2. The sensitivity of the volatility surface (in state 2) to parameters is studied in the

next two subplots of figure 9. Increasing η or increasing α respectively shifts up and down the volatility surface by a few percents. The last figure emphasizes that the asymmetry of the smile may be controlled by parameters ρ^+ and ρ^- , defining the double exponential distribution of jumps.

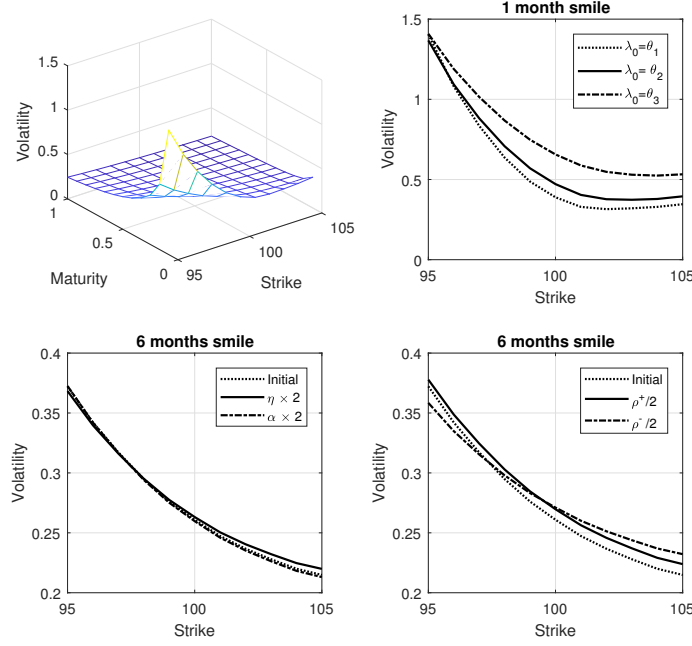


Figure 9: Smiles of volatilities for European call options. Maturities and strikes respectively range from 1 to 12 months and from 95 to 105. The risk free rate is set to 1%, and parameters used are these obtained with the PMCMC algorithm.

Tables 13 and 14 disentangle the influence of jump clustering and regime switches on the smile of implied volatility of call options, in the second and third economic regimes. Removing the jump clustering feature reduces the implied volatility but the amplitude depends upon the initial economic regime and upon the expiry of the option. In regime 3, the frequency of jumps is higher and the part of implied volatility related to jump clustering is nearly four times higher than in regime 2. However, the spread of implied volatility caused by jump clustering clearly decreases with time because the probability of transiting to the first regime increases. The regime-switching feature has a bigger impact on the implied volatility than jump clustering. Assuming that the economy stays blocked in the second regime slightly increases at short term and raises at medium term the implied volatility. If the economy does not leave the third regime, implied volatilities increase by several percent, mainly because the Brownian volatility is much higher in this situation.

K=100	Implied Vol.	impact of jump clustering	impact of regime switches
3 M	33.26%	-0.23%	+0.73%
6 M	26.08%	-0.16%	+0.45%
9 M	22.38%	-0.11%	-0.05%
12 M	19.82%	-0.10%	-0.35%
K=105	Implied Vol.	impact of jump clustering	impact of regime switches
3 M	25.54%	-0.59%	+0.78%
6 M	21.48%	-0.29%	+0.26%
9 M	19.37%	-0.21%	-0.18%
12 M	17.70%	-0.16%	-0.41%

Table 13: Analysis of the impact of regime shifts and jump clustering on the implied volatility of call options. Assumption: $\delta(t) = e_2$, $\lambda_0 = \theta_2$, $S_0 = 100$, strike $K = 100$ and $K = 105$.

K=100	Implied Vol.	impact of jump clustering	impact of regime switches
3 M	39.08%	-0.89%	+5.53%
6 M	28.78%	-0.52%	+2.17%
9 M	24.20%	-0.55%	+0.24%
12 M	21.10%	-0.54%	-0.69%
K=105	Implied Vol.	impact of jump clustering	impact of regime switches
3 M	31.06%	-0.80%	+8.52%
6 M	23.75%	-0.53%	+5.66%
9 M	20.71%	-0.56%	+3.47%
12 M	18.57%	-0.49%	+2.24%

Table 14: Analysis of the impact of regime shifts and jump clustering on the implied volatility of call options. Assumption: $\delta(t) = e_2$, $\lambda_0 = \theta_2$, $S_0 = 100$, strike $K = 100$ and $K = 105$.

6 Conclusions

This work introduces the first switching self-excited jump diffusion (SSEJD) process for modelling equity returns. A hidden Markov chain, interpreted as the economic conjuncture, determines the drift, the Brownian volatility and the mean reversion level of jumps frequency involved in the dynamics of stocks. The SSEJD presents several interesting features. Firstly, it replicates the clustering of jumps observed in financial markets during period of recessions. Secondly, it takes into account economic cycles.

We provide in this framework closed and semi-closed form expressions for the mean-variance of the intensity and for the joint moment generating of log returns and intensity. A sequential Monte-Carlo is developed to filter hidden states variables and is next combined with a Monte-Carlo Markov Chain algorithm to estimate the SSEJD. A POT method is developed to initialize the PMCMC procedure.

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 the economic conjuncture. The analysis also emphasizes that self-excited jumps are absent during periods of stable economic growth, as e.g. between 2004 and 2007. Whereas, most of the volatility of the S&P 500 is explained by the jump process during periods of economic turmoil. Our empirical study also emphasizes that in non switching self excited models, the reversion level of the intensity is underestimated during economic recessions.

Our model may also be used to price options. We show that Esscher changes of measure preserve the dynamics of the process under the risk neutral measure. As the moment generating function of the SSJED can be inverted numerically to retrieve the probability density function of log-prices, the evaluation of European options does not present any particular difficulty.

The Brownian volatility of the SSEJD model is stochastic but takes a limited number of values. This observation sketches some paths for future research and new challenges. A natural extension of the SSEJD consists to replace the switching Brownian variance by a switching mean-reverting process with self-excited jumps proportional to jumps of prices. The process obtained by this way would be able to explain both the clustering of jumps and the clustering of volatility. The calibration of such a process can only be done by a PMCMC algorithm. In this case, it is not anymore possible to find a good prior distribution for parameters with a POT procedure.

Appendix A: proofs

Proof of proposition 2.1. As $\mathcal{F}_0 \subset \mathcal{F}_0 \vee \mathcal{G}_t$, the expectation of λ_t is nested as follows:

$$\mathbb{E}(\lambda_t | \mathcal{F}_0) = \mathbb{E}(\mathbb{E}(\lambda_t | \mathcal{F}_0 \vee \mathcal{G}_t) | \mathcal{F}_0)$$

and if we remember the expression (12) for the intensity, we infer that

$$\mathbb{E}(\lambda_t | \mathcal{F}_0 \vee \mathcal{G}_t) = \lambda_0 - \alpha \int_0^t e^{\alpha(s-t)} (\lambda_0 - \theta_s) ds + \int_0^t \eta e^{\alpha(s-t)} \mathbb{E}(dL_s | \mathcal{F}_0 \vee \mathcal{G}_t) .$$

Let us denote by $\chi(t, l)$, the random measure of $L_t | \mathcal{F}_0 \vee \mathcal{G}_t$. $\chi(t, l)$ is such that $L_t | \mathcal{F}_0 \vee \mathcal{G}_t = \int_0^t \int_0^\infty \chi(ds, dl)$ and $d(L_t | \mathcal{F}_0 \vee \mathcal{G}_t) = \int_0^\infty \chi(ds, dl)$. Given that the integrand in the last term of equation (??) is a function of time integrable on \mathbb{R}^+ , independent from the random measure χ , we infer that

$$\begin{aligned} \mathbb{E} \left(\int_0^t \eta e^{\alpha(s-t)} dL_s | \mathcal{F}_0 \vee \mathcal{G}_t \right) &= \int_0^t \int_0^\infty \eta e^{\alpha(s-t)} \chi(ds, dl) \\ &= \int_0^t \eta e^{\alpha(s-t)} \int_0^\infty \chi(ds, dl) \\ &= \int_0^t \eta e^{\alpha(s-t)} \mathbb{E}(dL_s | \mathcal{F}_0 \vee \mathcal{G}_t) . \end{aligned}$$

According to the Itô's lemma for semi-martingale and by construction of L_s , L_s is solution of the following SDE:

$$dL_s = \int_0^\infty (L_s + |z| - L_s) \chi(ds, dl) = |J| dN_s .$$

Given that the jump J is independent from N_t , the expectation of dL_s with respect to $\mathcal{F}_0 \vee \mathcal{G}_t$ is then

$$\mathbb{E}(dL_s | \mathcal{F}_0 \vee \mathcal{G}_t) = \mathbb{E}(|J|) \times \mathbb{E}(dN_s | \mathcal{F}_0 \vee \mathcal{G}_s) \quad (36)$$

for $s \leq t$. Furthermore, conditionally to the filtration of λ_t denoted $(\mathcal{L}_t)_{t \geq 0}$, N_t is a Poisson random variable of intensity and expectation equal to $\int_0^{t-} \lambda_u du$. Using nested expectations allows us to infer that

$$\begin{aligned} \mathbb{E}(dN_s | \mathcal{F}_0 \vee \mathcal{G}_s) &= \mathbb{E}(\mathbb{E}(dN_s | \mathcal{F}_0 \vee \mathcal{G}_s \vee \mathcal{L}_s) | \mathcal{F}_0 \vee \mathcal{G}_s) \\ &= \mathbb{E}(\lambda_{s-} | \mathcal{F}_0 \vee \mathcal{G}_s) . \end{aligned} \quad (37)$$

Combining (36) and (37) leads then to

$$\mathbb{E}(dL_s | \mathcal{F}_0 \vee \mathcal{G}_t) = \mathbb{E}(|J|) \times \mathbb{E}(\lambda_{s-} | \mathcal{F}_0 \vee \mathcal{G}_s) ds$$

and

$$\mathbb{E}(\lambda_t | \mathcal{F}_0 \vee \mathcal{G}_t) = \lambda_0 - \alpha \int_0^t e^{\alpha(s-t)} (\lambda_0 - \theta_s) ds + \eta \mathbb{E}(|J|) \int_0^t e^{\alpha(s-t)} \mathbb{E}(\lambda_{s-} | \mathcal{F}_0 \vee \mathcal{G}_s) ds$$

If we derive this last expression with respect to time, we find that $\mathbb{E}(\lambda_t | \mathcal{F}_0 \vee \mathcal{G}_t)$ is solution of an ordinary differential equation (ODE):

$$\begin{aligned} \frac{\partial}{\partial t} \mathbb{E}(\lambda_t | \mathcal{F}_0 \vee \mathcal{G}_t) &= -\alpha (\lambda_0 - \theta_t) + \alpha^2 \int_0^t e^{\alpha(s-t)} (\lambda_0 - \theta_s) ds + \eta \mathbb{E}(|J|) \mathbb{E}(\lambda_t | \mathcal{F}_0 \vee \mathcal{G}_t) \\ &\quad - \alpha \eta \mathbb{E}(|J|) \int_0^t e^{-\alpha(t-s)} \mathbb{E}(\lambda_{s-} | \mathcal{F}_0 \vee \mathcal{G}_s) ds \\ &= (\eta \mathbb{E}(|J|) - \alpha) \mathbb{E}(\lambda_t | \mathcal{F}_0 \vee \mathcal{G}_t) + \alpha \theta_t . \end{aligned} \quad (38)$$

The solution of this ODE is the following function:

$$\mathbb{E}(\lambda_t | \mathcal{F}_0 \vee \mathcal{G}_t) = \int_0^t \alpha \theta_s e^{(\eta \mathbb{E}(|J|) - \alpha)(t-s)} ds + \lambda_0 e^{(\eta \mathbb{E}(|J|) - \alpha)t} ,$$

and the \mathcal{F}_0 conditional expectation is given by:

$$\begin{aligned} \mathbb{E}(\mathbb{E}(\lambda_t | \mathcal{F}_0 \vee \mathcal{G}_t) | \mathcal{F}_0) &= \mathbb{E} \left(\int_0^t \alpha \theta_s e^{(\eta \mathbb{E}(|J|) - \alpha)(t-s)} ds + \lambda_0 e^{(\eta \mathbb{E}(|J|) - \alpha)t} | \mathcal{F}_0 \right) \\ &= \alpha \int_0^t e^{(\eta \mathbb{E}(|J|) - \alpha)(t-s)} \mathbb{E}(\theta_s | \mathcal{F}_0) ds + \lambda_0 e^{(\eta \mathbb{E}(|J|) - \alpha)t} . \end{aligned} \quad (39)$$

The expected level of mean reversion at time s is the sum of $\bar{\theta}_j$ for $j = 1$ to N weighted by the probabilities of transition:

$$\mathbb{E}(\theta_s | \mathcal{F}_0) = \delta(0)' \exp(Q_0 s) \bar{\theta}$$

and if I is the $N \times N$ identity matrix, the integral in (39) may be rewritten as follows:

$$\begin{aligned} \int_0^t e^{(\eta \mathbb{E}(|J|) - \alpha)(t-s)} \mathbb{E}(\theta_s | \mathcal{F}_0) ds &= e^{(\eta \mathbb{E}(|J|) - \alpha)t} \int_0^t e^{-(\eta \mathbb{E}(|J|) - \alpha)s} \delta(0)' \exp(Q_0 s) \bar{\theta} ds \\ &= e^{(\eta \mathbb{E}(|J|) - \alpha)t} \int_0^t \delta(0)' \exp((Q_0 - I(\eta \mathbb{E}(|J|) - \alpha))s) \bar{\theta} ds \\ &= e^{(\eta \mathbb{E}(|J|) - \alpha)t} \left[\delta(0)' (Q_0 - I(\eta \mathbb{E}(|J|) - \alpha))^{-1} \exp((Q_0 - I(\eta \mathbb{E}(|J|) - \alpha))s) \bar{\theta} \right]_{s=0}^{s=t} \\ &= \delta(0)' (Q_0 - I(\eta \mathbb{E}(|J|) - \alpha))^{-1} [\exp(Q_0 t) - \exp(I(\eta \mathbb{E}(|J|) - \alpha)t)] \bar{\theta} \end{aligned}$$

■

Proof of Corollary 2.2. Conditionally to \mathcal{G}_t , this expectation is equal to

$$\begin{aligned}\mathbb{E}(\theta_t \lambda_t | \mathcal{F}_0 \vee \mathcal{G}_t) &= \theta_t \mathbb{E}(\lambda_t | \mathcal{F}_0 \vee \mathcal{G}_t) \\ &= \theta_t \int_0^t \alpha \theta_s e^{(\eta \mathbb{E}(|J|) - \alpha)(t-s)} ds + \lambda_0 \theta_t e^{(\eta \mathbb{E}(|J|) - \alpha)t}.\end{aligned}$$

We infer from this relation that the derivative with respect to time is given by

$$\begin{aligned}\frac{\partial}{\partial t} \mathbb{E}(\theta_t \lambda_t | \mathcal{F}_0 \vee \mathcal{G}_t) &= \theta_t \frac{\partial}{\partial t} \mathbb{E}(\lambda_t | \mathcal{F}_0 \vee \mathcal{G}_t) \\ &= (\eta \mathbb{E}(|J|) - \alpha) \mathbb{E}(\theta_t \lambda_t | \mathcal{F}_0 \vee \mathcal{G}_t) + \alpha (\theta_t)^2.\end{aligned}$$

and we can conclude that $\mathbb{E}(\theta_t \lambda_t | \mathcal{F}_0)$ is well given by equation (15). ■

Proof of proposition 2.3. Let us introduce the following notations: $f(t, N_t, \lambda_t, \delta_t) = \lambda_t^2$, according to the Itô's lemma for semi-martingale, the infinitesimal generator of this function is given by:

$$\begin{aligned}\mathcal{A}f(t, N_t, \lambda_t, \delta_t) &= \alpha(\theta_t - \lambda_t) 2\lambda_t + \lambda_t \int_{-\infty}^{+\infty} (\lambda_t + \eta |z|)^2 - \lambda_t^2 d\nu(z) \\ &= \alpha(\theta_t - \lambda_t) 2\lambda_t + \lambda_t \int_{-\infty}^{+\infty} (\lambda_t^2 + 2\eta |z| \lambda_t + \eta^2 |z|^2) - \lambda_t^2 d\nu(z) \\ &= 2\alpha \theta_t \lambda_t + 2(\eta \mathbb{E}(|J|) - \alpha) \lambda_t^2 + \eta^2 \lambda_t \mathbb{E}(|J|^2)\end{aligned}$$

On the other hand, we can prove that

$$\begin{aligned}\mathbb{E}(f(t, N_t, \lambda_t, \delta_t) | \mathcal{F}_0) &= f(0, N_0, \lambda_0, \delta_0) + \mathbb{E}\left(\int_0^t \mathcal{A}f(s, N_s, \lambda_s, \delta_s) ds | \mathcal{F}_0\right) \\ &= f(0, N_0, \lambda_0, \delta_0) + \int_0^t \mathbb{E}(\mathcal{A}f(s, N_s, \lambda_s, \delta_s) | \mathcal{F}_0) ds.\end{aligned}\tag{40}$$

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

$$\frac{\partial}{\partial t} \mathbb{E}(f(t, N_t, \lambda_t, \delta_t) | \mathcal{F}_0) = \mathbb{E}(\mathcal{A}f(t, N_t, \lambda_t, \theta_t) | \mathcal{F}_0),\tag{41}$$

that allows us to infer equation (16). It is easy to check that the right term of equation (41) is linear and Lipschitz. Then, according to the theorem of Cauchy-Lipschitz the solution exists and is unique. ■

Proof of proposition 2.4. If we note $f(t, N_t, \lambda_t, \delta_t) = \mathbb{E}(\omega^{N_s} | \mathcal{F}_t)$, f is solution of an Itô's equation for semi martingale. If $\delta_t = e_i$ then:

$$\begin{aligned}0 &= f_t + \alpha(\bar{\theta}_i - \lambda_t) f_\lambda + \lambda_t \int_{-\infty}^{+\infty} f(t, N_t + 1, \lambda_t + \eta |z|, e_i) - f(\cdot) d\nu(z) \\ &\quad + \sum_{j \neq i}^N q_{i,j} (f(t, N_t, \lambda_t, e_j) - f(t, N_t, \lambda_t, e_i))\end{aligned}\tag{42}$$

If we assume that f is an exponential affine function of λ_t and N_t :

$$f = \exp(A(t, s, e_i) + B(t, s)\lambda_t + C(t, s)N_t),$$

where $A(t, s, e_i)$ for $i = 1$ to N , $B(t, s)$, $C(t, s)$ are time dependent functions, the partial derivatives of f are given by:

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

$$f_\lambda = B(t, s)f$$

The integral in equation (42) is rewritten as follows:

$$\begin{aligned} & \int_{-\infty}^{+\infty} f(t, N_t + 1, \lambda_t + \eta|z|, e_j) - f(t, N_t, \lambda_t, e_i) d\nu(z) \\ &= \int_0^{+\infty} \exp(A(t, s, e_j) + B(t, s)\lambda_t + C(t, s)N_t) \left(e^{B(t, s)\eta|z| + C(t, s)} - 1 \right) d\nu(z) \\ &= f[e^C \psi(0, B(t, s)\eta) - 1]. \end{aligned}$$

As $q_{ii} = -\sum_{j \neq i}^N q_{i,j}$, the last term of equation (42) becomes

$$\sum_{j \neq i}^N q_{i,j} (f(t, N_t, \lambda_t, e_j) - f(t, N_t, \lambda_t, e_i)) = \sum_{j=1}^N q_{i,j} f(t, N_t, \lambda_t, e_j).$$

Then

$$\begin{aligned} 0 &= \left(\frac{\partial}{\partial t} A + \frac{\partial}{\partial t} B \lambda_t + \frac{\partial}{\partial t} C N_t \right) e^{A(t, s, e_i)} + \alpha(\bar{\theta}_i - \lambda_t) B e^{A(t, s, e_i)} \\ &+ \lambda_t [e^C \psi(0, B(t, s)\eta) - 1] e^{A(t, s, e_i)} + \sum_{j=1}^N q_{i,j} e^{A(t, s, e_j)}, \end{aligned}$$

from which we deduce that $C(t, s) = \ln \omega$. Regrouping terms allows to infer that

$$\begin{aligned} 0 &= \left(\frac{\partial}{\partial t} A \right) e^{A(t, s, e_i)} + \alpha \bar{\theta}_i B e^{A(t, s, e_i)} + \sum_{j=1}^N q_{i,j} e^{A(t, s, e_j)} \\ &+ \lambda_t \left(\frac{\partial}{\partial t} B - \alpha B + [\omega \psi(0, B(t, s)\eta) - 1] \right) e^{A(t, s, e_i)} \end{aligned}$$

or that

$$\begin{aligned} \frac{\partial}{\partial t} B &= \alpha B - [\omega \psi(0, B(t, s)\eta) - 1] \\ \left(\frac{\partial}{\partial t} A \right) e^{A(t, s, e_i)} &= -\alpha \bar{\theta}_i B e^{A(t, s, e_i)} - \sum_{j=1}^N q_{i,j} e^{A(t, s, e_j)} \end{aligned}$$

If we define $\tilde{A}(t, s, e_i) = e^{A(t, s, e_i)}$, the first equations can finally be put in matrix form as:

$$\frac{\partial \tilde{A}(t)}{\partial t} + (\text{diag}(\alpha \bar{\theta} B) + Q_0) \tilde{A}(t) = 0, \quad (43)$$

■

Proof of proposition 2.5. If we note $f(t, X_t, \lambda_t, \delta_t) = \mathbb{E}(e^{\omega_1 X_s + \omega_2 \lambda_s} | \mathcal{F}_t)$, f is solution of an Itô's equation for semi-martingale. If $\delta_t = e_i$ then:

$$\begin{aligned} 0 &= f_t + f_X \left(\bar{\mu}_i - \frac{\bar{\sigma}_i^2}{2} - \lambda_t \mathbb{E}(e^J - 1) \right) + f_{XX} \frac{\bar{\sigma}_i^2}{2} + \alpha(\bar{\theta}_i - \lambda_t) f_\lambda + \\ &\lambda_t \int_{-\infty}^{+\infty} f(t, X_t + z, \lambda_t + \eta|z|, e_i) - f(\cdot) \nu(dz) + \sum_{j \neq i}^N q_{i,j} (f(t, X_t, \lambda_t, e_j) - f(t, X_t, \lambda_t, e_i)) \end{aligned} \quad (44)$$

If we assume that f is an exponential affine function of λ_t and X_t :

$$f = \exp(A(t, s, e_i) + B(t, s)\lambda_t + C(t, s)X_t),$$

where $A(t, s, e_i)$ for $i = 1$ to N , $B(t, s)$, $C(t, s)$ are time dependent functions, the partial derivatives of f are given by:

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

$$f_X = C(t, s)f \quad f_{XX} = C(t, s)^2 f \quad f_\lambda = B(t, s)f$$

and the integrand in equation (44) is rewritten as follows:

$$\int_{-\infty}^{+\infty} f(t, X_t + z, \lambda_t + \eta|z|, e_i) - f(\cdot) \nu(dz) = f[\psi(C(t, s), B(t, s)\eta) - 1].$$

As $q_{ii} = -\sum_{i \neq j}^N q_{i,j}$, the last term of the Itô equation is also equal to

$$\sum_{j \neq i}^N q_{i,j} (f(t, X_t, \lambda_t, e_j) - f(t, X_t, \lambda_t, e_i)) = \sum_{j=1}^N q_{i,j} f(t, X_t, \lambda_t, e_j)$$

injecting these expressions into equation (44), leads to the following relation:

$$\begin{aligned} 0 = & \left(\frac{\partial}{\partial t} A + \frac{\partial}{\partial t} B \lambda_t + \frac{\partial}{\partial t} C X_t \right) e^{A(t,s,e_i)} + C \left(\bar{\mu}_i - \frac{\bar{\sigma}_i^2}{2} - \lambda_t \mathbb{E}(e^J - 1) \right) e^{A(t,s,e_i)} \\ & + C^2 \frac{\bar{\sigma}_i^2}{2} e^{A(t,s,e_i)} + \alpha(\bar{\theta}_i - \lambda_t) B e^{A(t,s,e_i)} + \lambda_t [\psi(C, B\eta) - 1] e^{A(t,s,e_i)} + \sum_{j=1}^N q_{i,j} e^{A(t,s,e_j)}, \end{aligned}$$

from which we deduce that $C(t, s) = \omega_1$. Regrouping terms allows to infer that A and B are solutions of a system of ODE's:

$$\begin{aligned} 0 = & \left(\frac{\partial}{\partial t} A + \frac{\partial}{\partial t} B \lambda_t \right) e^{A(t,s,e_i)} + \omega_1 \left(\bar{\mu}_i - \frac{\bar{\sigma}_i^2}{2} - \lambda_t \mathbb{E}(e^J - 1) \right) e^{A(t,s,e_i)} \\ & + \omega_1^2 \frac{\bar{\sigma}_i^2}{2} e^{A(t,s,e_i)} + \alpha(\bar{\theta}_i - \lambda_t) B e^{A(t,s,e_i)} + \lambda_t [\psi(\omega_1, B\eta) - 1] e^{A(t,s,e_i)} + \sum_{j=1}^N q_{i,j} e^{A(t,s,e_j)}, \end{aligned}$$

From this last relation, we deduce that

$$\begin{cases} 0 = \frac{\partial}{\partial t} A e^{A(t,s,e_i)} + \omega_1 \left(\bar{\mu}_i - \frac{\bar{\sigma}_i^2}{2} \right) e^{A(t,s,e_i)} + \omega_1^2 \frac{\bar{\sigma}_i^2}{2} e^{A(t,s,e_i)} \\ \quad + \alpha \bar{\theta}_i B e^{A(t,s,e_i)} + \sum_{j=1}^N q_{i,j} e^{A(t,s,e_j)} \text{ for } i = 1 \dots N, \\ 0 = \frac{\partial}{\partial t} B - \alpha B - \omega_1 \mathbb{E}(e^J - 1) + [\psi(\omega_1, B\eta) - 1] \end{cases}$$

If we define $\tilde{A}(t, s) = (e^{A(t,s,e_i)})_{i=1, \dots, N}$, the first equations can finally be put in matrix form as:

$$\frac{\partial \tilde{A}(t, s)}{\partial t} + \left(\text{diag} \left(\omega_1 \left(\bar{\mu} - \frac{\bar{\sigma}^2}{2} \right) + \omega_1^2 \frac{\bar{\sigma}^2}{2} + \alpha \bar{\theta} B \right) + Q_0 \right) \tilde{A}(t, s) = 0, \quad (45)$$

■

Proof of proposition From the previous results, we know that $B(t, s)$ is solution of the ODE:

$$\frac{\partial}{\partial t} B = \alpha B + \omega_1 (\psi(1, 0) - 1) - [\psi(\omega_1, B\eta) - 1]$$

with the terminal condition: $B(s, s) = \omega_2$. If we set $B(t, s) = C(s - t)$ and $\tau = s - t$. Then $\frac{\partial}{\partial t} B = \frac{\partial}{\partial \tau} C \frac{\partial \tau}{\partial t} = -\frac{\partial}{\partial \tau} C$ and

$$\begin{aligned} \frac{\partial}{\partial \tau} C(\tau) &= -\alpha C(\tau) + \psi(\omega_1, C(\tau)\eta) - \underbrace{[\omega_1 (\psi(1, 0) - 1) + 1]}_{\beta(\omega_1)} \\ &= -\alpha C(\tau) + \psi(\omega_1, C(\tau)\eta) - \beta(\omega_1) \end{aligned} \quad (46)$$

The left hand side is denoted $h(C)$. Due to the convexity of $\psi(\cdot)$ there is only one point u^* such that $h(u) = 0$. This equation is indeed equivalent to

$$\psi(\omega_1, u\eta) = \beta(\omega_1) + \alpha u$$

We rewrite the equation (46) as follows,

$$\frac{dC}{-\beta - \alpha C + \psi(\omega_1, C\eta)} = d\tau.$$

As $C(0) = \omega_2$, by direct integration, we have that

$$\int_{\omega_2}^C \frac{du}{-\beta - \alpha u + \psi(\omega_1, \eta u)} = \tau$$

with $C \in [\omega_2, u^*)$ or $C \in [u^*, \omega_2)$. Remark that if $C = u^*$ then $\tau = +\infty$ as the numerator converges to zero. If we define the function on the left hand side as

$$F_{\omega_1}(C) := \int_{\omega_2}^C \frac{du}{-\beta - \alpha u + \psi(\omega_1, \eta u)}$$

then $F_{\omega_1}(C) = \tau$ and $C = F_{\omega_1}^{-1}(\tau)$ or $B(t) = F_{\omega_1}^{-1}(s - t)$. ■

Appendix B: particle filter

The structure of the particle filter algorithm is the following:

1. **Initial step:** draw M values of $v_0^{(i)}$ for $i = 1, \dots, M$, from an initial distribution $p(v_0)$
2. For $j = 1 : T$
Prediction step: draw a sample of $\Delta L_j^{(i)}$ and $\delta_j^{(i)}$ and update $\lambda_j^{(i)}, \mu_j^{b(i)} \sigma_j^{(i)}, \theta_j^{(i)}$ using the relations

$$\begin{aligned} \lambda_j^{(i)} &= \lambda_{j-1}^{(i)} + \alpha(\theta_{j-1}^{(i)} - \lambda_{j-1}^{(i)})\Delta + \eta \Delta L_j^{(i)} \\ \mu_j^{(i)} &= \delta_j^{(i)} \bar{\mu} \quad \sigma_j^{(i)} = \delta_j^{(i)} \bar{\sigma} \quad \theta_j^{(i)} = \delta_j^{(i)} \bar{\theta} \end{aligned}$$

Correction step: the particle $v_j^{(i)}$ has a probability of occurrence equal to $w_j^{(i)} = \frac{p(x_j | v_j)}{\sum_{i=1:M} p(x_j | v_j)}$ where

$$p(x_j | v_j^{(i)}) \sim \mathcal{N} \left(\left(\mu_j^{(i)} - \frac{\sigma_j^{(i)2}}{2} - \lambda_j^{(i)} \mathbb{E}(e^J - 1) \right) \Delta - \Delta L_j^{(i)}, \sigma_j^{(i)} \sqrt{\Delta} \right)$$

Resampling step: resample with replacement M particles according to the importance weights $w_j^{(i)}$. The new importance weights are set to $w_j^{(i)} = \frac{1}{M}$.

References

- [1] Alexander, S.S., 1964, Price Movements in Speculative Markets: Trends or Random Walks. in *Cootner* (1964), 338-372
- [2] Andersen, T., Benzoni, L., Lund, J., 2002. An empirical investigation of continuous-time equity return models. *Journal of Finance* 57, 1239-1284.
- [3] 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.
- [4] Ait-Sahalia, R. J. A. Laeven, and L. Pelizzon. 2014. Mutual excitation in eurozone sovereign CDS. *Journal of Econometrics*, 183(2), 151–167.
- [5] Al-Anaswah N., Wilfing B. 2011. Identification of speculative bubbles using state-space models with Markov-switching. *Journal of Banking & Finance*, 35(5), 1073-1086.
- [6] Bates, D., 2000. Post-87 crash fears in S&P500 futures options. *Journal of Econometrics* 94, 181-238.
- [7] Bollerslev T., 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31, 307-27.
- [8] Calvet L., Fisher A. 2001. Forecasting multifractal volatility. *Journal of econometrics*. 105, 17-58.
- [9] Calvet L., Fisher A. 2004. How to forecast long-Run volatility: regime switching and the estimation of multifractal processes. *Journal of Financial Economics* 2, 49-83.
- [10] Carr P., Wu L. 2016. Leverage Effect, Volatility Feedback, and Self-Exciting Market Disruptions. *Forthcoming in Journal of Financial and Quantitative Analysis*.
- [11] Chen K., Poon S.-H. 2013. Variance swap premium under stochastic volatility and self-exciting jumps. *working paper SSRN-id2200172*.
- [12] Cholette L., Heinen A., Valdesogo Al. 2009. Modelling international financial returns with a multivariate regime switching copula. *Journal of financial econometrics* 7 (4), 437-480.
- [13] Chopin, N., Jacob, P. E., and Papaspiliopoulos, O. 2013. SMC2: an efficient algorithm for sequential analysis of state space models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* 75(3) , 397-426.
- [14] Cont R. Tankov P. 2004. Financial modelling with jump processes. Chapman & Hall/CRC financial mathematics series.
- [15] Doucet, A., J.F.G. De Freitas and N. Gordon. *Sequential Monte Carlo Methods in Practice*. Cambridge University Press, Cambridge 2000.
- [16] Duffie D., Singleton K., 1993, Simulated moments estimation of Markov models of asset prices, *Econometrica* 61, 929–952.
- [17] Duffie D., Pan J., Singleton K. 2000. Transform analysis and asset pricing for affine jump diffusions, *Econometrica* 68, 1343-1376.
- [18] Elliott R. J., Chan L., Siu T. K. 2005. Option pricing and Esscher transform under regime switching. *Annals of Finance* 1, 423-432.
- [19] Elliott R.J., Siu T.K., Chan L. and Lau J.W. 2007. Pricing options under a generalized Markov-modulated jump-diffusion Model. *Stochastic Analysis and Applications*. 25(4), 821-843.

- [20] Embrechts P., Liniger T., Lu L., 2011, Multivariate Hawkes processes: an application to financial data. *Journal of Applied Probability*, 48 (A), 367–378.
- [21] Engel R., 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50, 987-1007.
- [22] Eraker, B., Johannes, M., Polson, N., 2003. The impact of jumps in volatility and returns. *Journal of Finance* 58, 1269-1300.
- [23] Evans K. 2011. Intraday jumps and US macroeconomic news announcements. *Journal of Banking & Finance* 35 (10), 2511-2527
- [24] Fulop A., Li J. 2013. Efficient learning via simulation: A marginalized resample-move approach. *Journal of Econometrics*, 176(2), 146-161
- [25] Fulop A., Duan J.-C. 2015. Density-tempered marginalized sequential Monte Carlo samplers. *Journal of Business and Economic Statistics*, 33, 192-202
- [26] Gallant, A. R., Tauchen G., 1996, Which moments to match? *Econometric Theory* 12, 657–681.
- [27] Gouriéroux C., Monfort A., Renault E., 1993, Indirect inference, *Journal of Applied Econometrics* 8, S85–S118
- [28] Fulop A., Li J., Yu J. 2015. Self-Exciting Jumps, Learning, and Asset Pricing Implications. *Review of Financial Studies* 28, 876-912.
- [29] Gerber H.U., Shiu E.S.W. 1994. Options Pricing by Esscher Transform. *Transactions of the society of actuaries* 26, 99-191.
- [30] Guidolin M., Timmermann A. 2005. Economic Implications of bull and bear regimes in UK stock and bond returns. *The Economic Journal* 115, 11-143
- [31] Guidolin M., Timmermann A. 2007. Asset allocation under multivariate regime switching. *Journal of Economic Dynamics and Control* 31 (11), 3503-3544.
- [32] Guidolin M., Timmermann A. 2008. International Asset Allocation under Regime Switching, Skew, and Kurtosis Preferences. *Review of Financial Studies* 21 (2), 889-935.
- [33] Hainaut D. 2016. A model for interest rates with clustering effects. *Quantitative finance* 16 (8), 1203-1218.
- [34] Hamilton J.D. 1989 .A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica* 57 (2), 357-384.
- [35] Hardy M. 2001. A Regime-Switching Model of Long-Term Stock Returns. *North American Actuarial Journal* 5(2), 41-53.
- [36] Hawkes A., 1971(a). Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society Series B*, 33, 438-443.
- [37] Hawkes A., 1971(b). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* 58, 83–90.
- [38] Hawkes A. and Oakes D., 1974. A cluster representation of a self-exciting process. *Journal of Applied Probability* 11, 493-503.

- [39] Horst U., Paulsen M. 2017. A Law of Large Numbers for Limit Order Books. *Mathematics of operations research*, Online : <https://doi.org/10.1287/moor.2017.0848>
- [40] Hunt L.H., Kavesh. A.R. 1976. Dynamics of forecasting financial cycles: Theory, technique and implementation. JAI Press.
- [41] Hunt L.H. 1987. A Time to be Rich: winning on Wall Street in the new economy. Rawson Associates, New York.
- [42] Kalman R.E.: A new approach to linear filtering and prediction problems. *J. Basic. Eng.* 82(1), 35-45 (1960).
- [43] Kelly F., Yudovina E., 2017. A Markov Model of a Limit Order Book: Thresholds, Recurrence, and Trading Strategies. *Mathematics of operations research*, Online : <https://doi.org/10.1287/moor.2017.0857>
- [44] Kyle A. 1985. Continuous auctions and insider trading. *Econometrica* 53, 1315-1335.
- [45] Lee S. Mykland P. 2008. Jumps in financial markets: a new nonparametric test and jump dynamics. *Review of Financial Studies* 21 (6), 2535-2563.
- [46] Maheu J., Chan W. 2002. Conditional Jump Dynamics in Stock Market Returns, *Journal of Business & Economic Statistics* 20(3), 377-389
- [47] Maheu, J., McCurdy, T., 2004. News arrival, jump dynamics and volatility components for individual stock returns. *Journal of Finance* 59, 755-793
- [48] Mandelbrot B., 1963. The Variation of Certain Speculative Prices, *Journal of Business*, 36, 394-419.
- [49] Nakajima J. 2013. Stochastic volatility model with regime-switching skewness in heavy-tailed errors for exchange rate returns. *Studies in Nonlinear Dynamics & Econometrics* 17 (5), p 499-520.
- [50] Osborne M.F. M., 1959. Brownian Motion in the Stock Market. *Operations Research*. 7, 145-173.
- [51] Schoutens W. 2003. *Lévy processes in Finance. Pricing of financial derivatives*. Eds Wiley.
- [52] Siu T.K. 2016. A self-exciting threshold jump-diffusion model for option valuation. *Insurance: Mathematics and Economics*. 69, 168-193.
- [53] Stovall S. 1996. Standard & Poor's Sector Investing: How to Buy The Right Stock in The Right Industry at The Right Time. McGraw-Hill Companies.
- [54] Wang T., Bebbington M., Harte D. 2012. Markov-modulated Hawkes process with stepwise decay. *Ann. Inst. Stat. Math.* 64, 521-544.