# Self-Exciting Jumps in the Oil Market: Bayesian Estimation and Dynamic Hedging

Luca Gonzato*, Carlo Sgarra†

**Abstract**

In this paper we propose a novel self-exciting jump-diffusion model for oil price dynamics based on a Hawkes-type process. In particular, the jump intensity is stochastic and path dependent, implying that the occurrence of a jump will increase the probability of observing a new jump and this feature of the model aims at explaining the jumps clustering effect. Moreover, volatility is described by a stochastic process, which can jump simultaneously with prices. The model specification is completed by a stochastic convenience yield. In order to estimate the model we apply the two-stage Sequential Monte Carlo (SMC) sampler (Fulop and Li, 2019) to both spot and futures quotations. From the estimation results we find evidence of self-excitation in the oil market, which leads to an improved fit and a better out of sample futures forecasting performance with respect to jump-diffusion models with constant intensity. Furthermore, we compute and discuss two optimal hedging strategies based on futures trading. The optimality of the first hedging strategy proposed is based on the variance minimization, while the second strategy takes into account also the third-order moment contribution in considering the investors attitudes. A comparison between the two strategies in terms of hedging effectiveness is provided.

Keywords: Hawkes Processes, Jumps Clustering, Oil Price Dynamics, Particle Filtering, Sequential Monte Carlo, Optimal Hedging.

*Dipartimento di Statistica e Metodi Quantitativi, Università degli Studi di Milano-Bicocca, Via Bicocca degli Arcimboldi, 8, 20126, Milano, Italy, email: l.gonzato@campus.unimib.it

†Corresponding Author, Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci, 32, 20133, Milano, Italy, email: carlo.sgarra@polimi.it

DOI 10.1016/j.eneco.2021.105279

# 1 Introduction

An accurate description of oil price dynamics is crucial for financial applications like risk management, portfolio allocation and derivatives pricing. In addition, the oil market has a strong impact on most aspects of economics in a wider sense; for example, from a macroeconomic perspective oil prices can affect the world GDP growth (Kilian and Figfusson, 2013), the efficiency of oil usage and energy consumption (Wang, 2013), and the term structure of interest rates (Ioannidis and Ka, 2018). This explains the huge amount of literature devoted to provide reliable and accurate methods for oil price dynamics calibration and forecast (Baumeister and Kilian, 2015). One of the most popular approaches for modeling commodities is represented by factor models, in which continuous time stochastic differential equations describe the *factors* moving the price dynamics. In particular, among the others we recall the two factor model proposed by Gibson and Schwartz (1990), where the spot price is described by a Geometric Brownian Motion (GBM) and the convenience yield by a Ornstein-Uhlenbeck (OU) process. Later, Ribeiro and Hodges (2004) proposed a multi-factor model, in which the convenience yield is driven by a Cox-Ingersoll-Ross (CIR) process. This ensures that the convenience yield does not take negative values and hence the model excludes arbitrage. None of these models accounts neither for stochastic volatility nor for jumps.

In order to provide a better description of oil prices dynamics, Larsson and Nossman (2011) introduced jumps in both stochastic volatility and returns. They analysed daily spot prices of WTI crude oil from 1989 to 2009, and exploited a Markov Chain Monte Carlo (MCMC) method for the model estimation. In contrast to the literature mentioned before, they did not consider mean reversion in the spot price and also the convenience yield was not introduced as a stochastic factor (they only considered a constant convenience yield in the risk-neutral specification of the model). A similar framework has been adopted also by Brooks and Prokopczuk (2013).

Another stochastic volatility model with jumps for oil prices can be found in two recent papers by Fileccia and Sgarra (2015, 2018), where the authors included also a stochastic convenience yield. In their paper information from futures prices is included and the model is estimated under both the historical and risk-neutral measure via Particle Markov Chain Monte Carlo (PMCMC).

In our paper we address the problem of describing oil prices dynamics by implementing a different modeling strategy. In particular, we want to investigate if the inclusion of self-exciting effects provides a better understanding of price movements. Self-exciting features have been

already systematically investigated in several asset classes. Fulop *et al.* (2015) estimate a model that considers co-jumps between prices and volatility and self-exciting jump clustering, on the S&P 500 index data from 1980 to 2012; Ait-Sahalia *et al.* (2015) model financial contagion with mutually exciting jump processes and Maneesoonthorn *et al.* (2016) extend the latter model by introducing self-exciting jumps in volatility in a univariate framework, i.e. no mutual excitation between different assets is considered. Finally, we mention the paper by Fulop and Li (2019), in which the authors propose a non-affine self-exciting jump diffusion model with stochastic volatility together with a new estimation method, namely the two-stage SMC sampler. Their methodology is applied on the S&P500 and variance swaps observations.

In order to detect self-exciting features in the oil price dynamics we estimate the parameters of a Hawkes-type jump-diffusion model by a particle filtering method. The data set consists of both spot and future quotations of WTI Cushing (Oklahoma) crude oil ranging from January 8, 2008 to December 31, 2018. The particle filtering methodology we apply is based on a hybrid particle filter with a two-stage density tempered Sequential Monte Carlo Method of the same kind of that proposed by Fulop and Li (2019). The model we propose in the present paper exhibits some similarities with the models mentioned above. In particular, our model is similar to the affine version of that considered by Fulop and Li (2019), where in addition we introduce another stochastic differential equation, which describes the evolution of the convenience yield as an OU process. However, an interesting feature of the present work is that we apply this kind of jump diffusion model to WTI crude oil spot and futures quotations. Indeed, there is in the literature some intuition behind the existence of this feature also in commodity markets. For example, Filimonov *et al.* (2014) fit a simple Hawkes process to high frequency data related to many different asset classes, including oil market. Ma *et al.* (2018) consider a plethora of realized range-based volatility models and document an increasing accuracy in futures price volatility forecasting when a Hawkes process is included in the model.

In the present modelling framework jumps with self-exciting features are included in the crude oil spot price dynamics. A remarkable property of the resulting jump-diffusion model is that it is affine, and it allows an explicit computation of the prices of futures contracts as functions of the model parameters, thus providing accurate estimation results. In this paper we address also the issue of hedging against oil price variations. We provide the computation of an optimal dynamic hedging strategy, where optimality refers to an objective function taking into account not only second order effects (variance), but also skewness. The hedging results highlight the

importance of high order effects with respect to the standard approach based only on variance minimization.

The paper is organized as follows. In Section 2 we present our self-exciting jump diffusion model and introduce the optimal dynamic hedging problem. In Section 3 we present our estimation method, while in Section 4 we describe the data set and the results obtained. In section 5 we discuss the empirical results of the hedging application based on minimization of variance and skewness. In section 6 we provide some concluding remarks.

## 2 The Model

In this section we introduce the most relevant features of Hawkes processes by following the exposition of Da Fonseca and Zaatour (2014).

The Hawkes process we shall consider throughout the paper exhibits an exponentially decaying intensity. This is a special kind of point process whose conditional intensity depends on the history of the events. Then, let us define in broader terms the conditional intensity function:

**Definition 1.** *Let $N_t$ be a point process and let $\mathcal{F}_t^N$ be the natural filtration generated by $N$ itself. Then, the left continuous process defined by:*

$$\lambda(t \mid \mathcal{F}_{t-}^N) = \lim_{h \to 0^+} \frac{\mathbb{P}\left[N_{t+h} - N_t > 0 \mid \mathcal{F}_{t-}^N\right]}{h} \tag{1}$$

*is called the conditional intensity function of the point process.*

Therefore, we have the following

**Definition 2.** *The univariate Hawkes process $N$ with conditional intensity $\lambda(t \mid \mathcal{F}_{t-}^N) = \lambda_t$, can be defined for all $t > 0$ and $h \to 0^+$ as:*

$$\begin{cases} \mathbb{P}\left[N_{t+h} - N_t = 1 \mid \mathcal{F}_{t-}^N\right] & = \lambda_t h + o(h) \\ \mathbb{P}\left[N_{t+h} - N_t > 1 \mid \mathcal{F}_{t-}^N\right] & = o(h) \\ \mathbb{P}\left[N_{t+h} - N_t = 0 \mid \mathcal{F}_{t-}^N\right] & = 1 - \lambda_t h + o(h) \end{cases} \tag{2}$$

In particular, we can represent the intensity of a Hawkes process by a stochastic differential equation (SDE):

$$d\lambda_t = \beta(\lambda_\infty - \lambda_t)dt + \alpha dN_t, \tag{3}$$

where, $\alpha \in \mathbb{R}^+$ is the magnitude of self-excited jump, $\beta \in \mathbb{R}^+$ is the constant rate of decay and $\lambda_\infty \in \mathbb{R}^+$ is the so called background intensity. Then, by applying Itô formula to $f(t, \lambda_t) = e^{\beta t}\lambda_t$, we obtain the solution to the above SDE:

$$\begin{aligned}\lambda_t &= \lambda_\infty + (\lambda_0 - \lambda_\infty)e^{-\beta t} + \int_0^t \alpha e^{-\beta(t-s)} dN_s, \\ &= \lambda_\infty + (\lambda_0 - \lambda_\infty)e^{-\beta t} + \sum_{j:t_j<t} \alpha e^{-\beta(t-t_j)}. \end{aligned} \tag{4}$$

The solution given by (4) allows to simulate the intensity process. Indeed, different simulation methods are proposed in the literature; for example, one of the most popular procedures is the Ogata's modified thinning algorithm (Ogata, 1981), which is based on the thinning method developed by Lewis and Shedler (1969) and further modified by Daley and Vere-Jones (2008). However, more recently Dassios and Zhao (2013) provided an "exact simulation" algorithm, which is faster than the previous approach. We can see an example of a simulated path of $\lambda_t$ in Figure 1, where the simulation is performed by applying the algorithm of Dassios and Zhao.

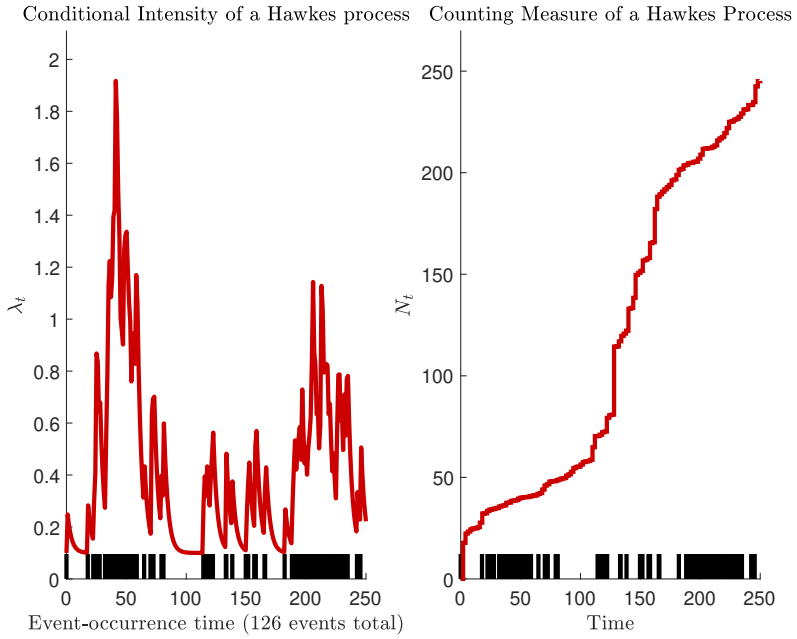


Figure 1: Simulated conditional intensity and counting measure of a Hawkes process on $[0, T]$ with parameters: $T = 250$, $\lambda_0 = \lambda_\infty = 0.1$, $\alpha = 0.2$ and $\beta = 0.3$.

From figure 1 we can clearly see the clustering of jumps exhibited by the Hawkes process; for

example, in the interval $[200, 240]$ the self-excitation effect looks quite relevant.

An important property of the Hawkes process under consideration is that, although $\lambda_t$ is clearly non-Markovian, it can be proved that the two-dimensional process $(N_t, \lambda_t)$ is jointly both a Markov and an affine process and this property improves significantly the analytical tractability of the model. In particular, in Da Fonseca and Zaatour (2014) the infinitesimal generator and the Dynkin's formula can be used in order to find some moments of the process as solutions of ordinary differential equations.

## 2.1 Model Dynamics Under The Historical Measure

Now, we can introduce our dynamic model, which takes into account co-jumps between stock price and volatility, stochastic convenience yield and self-excitation. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a complete filtration $(\mathcal{F}_t)_{t\geq 0}$, then our model for $X_t = \ln(S_t/S_0)$ is described by the following system of SDEs:

$$dX_t = \left(\mu - \frac{1}{2}V_t - \delta_t\right)dt + \sqrt{V_t}dW_t + dJ_{x,t}, \tag{5}$$

$$dV_t = k(\bar{V} - V_t)dt + \sigma_v\sqrt{V_t}dW_{v,t} + dJ_{v,t}, \tag{6}$$

$$d\delta_t = \gamma(\bar{\delta} - \delta_t)dt + \sigma_\delta\, dW_{\delta,t}, \tag{7}$$

$$d\lambda_t = \beta(\lambda_\infty - \lambda_t)dt + \alpha dN_t, \tag{8}$$

From equation (5) we see that changes in the underlying are given by a standard Brownian motion $W_t$ and compound Poisson process $J_{x,t}$, where the number of jumps $N_t$ is an Hawkes process with stochastic intensity $\lambda_t$. Furthermore, the amplitude of jumps is dictated by i.i.d Gaussian random variables with mean $\mu_J$ and variance $\sigma_J^2$. On the other hand, Equation (6) describes the evolution of the volatility, which is a mean reverting jump-diffusion process, where $W_{v,t}$ is a standard Brownian motion, possibly correlated with $W_t$, i.e. $\text{Corr}(dW, dW_v) = \rho_v dt$. This feature of the model is important in order to capture the so called *leverage effect*. Furthermore, recent studies find the presence of co-jumps of prices and volatility, not only on the equity market (Eraker, 2004, Eraker *et al.*, 2003, Fulop and Li, 2019), but also on the commodity market (Larsson and Nossman, 2011, Brooks and Prokopczuk, 2013). Moreover, there is evidence of jump clustering (Ait-Sahalia *et al.*, 2015, Fulop *et al.*, 2015), i.e. an extreme movement tends to be followed by another extreme movement. As a consequence, we introduce jumps in volatility

with $J_{v,t}$, which is a compound Poisson process with counting process $N_t$. To be more precise, returns and volatility jump together with the same self-exciting intensity $\lambda_t$ and the jump size of volatility follows an exponential distribution with mean $\mu_v$. Finally, since we are dealing with oil prices, there is an additional SDE, which describes the evolution of the convenience yield with a standard Ornstein-Uhlenbeck process as in Gibson and Schwartz (1990); Schwartz (1997); Lai and Mellios (2016); Yan (2002). The convenience yield includes both the reduction in cost of acquiring inventory and the value of being able to profit from temporary local shortage of the commodity (Yan, 2002), so it is natural to adopt a stochastic process, which could assume both positive and negative values.

In our empirical application we consider two nested models:

- Model I defined by Equations (5)-(8),
- Model II without self-exciting effect, i.e. $\beta = \alpha = 0$.

## 2.2 Risk-Neutral Dynamics and Futures Pricing

In order to perform joint estimation using spot and futures data, we need to derive a pricing formula for futures contracts. As usual, we employ a suitable change of measure form the real world measure $\mathbb{P}$ to the risk-neutral measure $\mathbb{Q}$. To this end consider the following Radon-Nikodym derivative:

$$\frac{d\mathbb{P}}{d\mathbb{Q}}\Big|_{\mathcal{F}_t} = \exp\left\{-\int_0^t r_s ds - \frac{1}{2}\int_0^t \varphi_x^2(u)du - \int_0^t \varphi_x(u)dW_u - \frac{1}{2}\int_0^t \varphi_\delta^2(u)du - \int_0^t \varphi_\delta(u)dW_u\right\}, \quad (9)$$

where, $r_t$ is the risk-free interest rate. Actually, as we shall see in next lines, we are pricing only the convenience yield risk, neither volatility nor jump risk. This is motivated by the fact the futures price cannot be a function of spot volatility, jumps or their associated parameters[1] (see Yan, 2002). However, volatility and jumps are important in order to describe appropriately spot prices and for hedging purpose, as we shall see later. Thus, by following the literature (e.g. Yu *et al.*, 2011, Pan, 2002, Fulop and Li, 2019) we left $\varphi_x(t)$ unspecified and choose the convenience

[1]The economic justification for this result relies in the linearity of the futures payoff, the local martingale behavior of the price with respect to the risk-neutral measure and the structure-preserving property of the measure change. Indeed, volatility and jumps affect higher order moments, but not the first one. From the mathematical point of view, the payoff of the contract is just the expected spot price under $\mathbb{Q}$; thus in the spot price dynamics we compensate the drift with the jump compensator $\lambda_t\mu^\star$ and the Itô term $1/2V_t$. As a consequence the solution of the respective ODEs is equal to zero.

yield risk premium as follows:

$$\varphi_\delta(t) = \frac{\varphi_\delta}{\sigma_\delta} \quad \Rightarrow \quad dW^{\mathbb{Q}}_{\delta,t} = dW^{\mathbb{P}}_{\delta,t} + \frac{\varphi_\delta}{\sigma_\delta} dt. \tag{10}$$

Thus, under the risk-neutral measure the structure of the model is preserved:

$$dX_t = \left( r - \frac{1}{2} V_t - \lambda_t \mu^\star - \delta_t \right) dt + \sqrt{V_t} dW_t + d\widetilde{J}_{x,t}, \tag{11}$$

$$dV_t = k(\bar{V} - V_t)dt + \sigma_v \sqrt{V_t}\, dW_{v,t} + dJ_{v,t}, \tag{12}$$

$$d\delta_t = \gamma(\bar{\delta}^{\mathbb{Q}} - \delta_t)dt + \sigma_\delta dW^{\mathbb{Q}}_{\delta,t}, \tag{13}$$

$$d\lambda_t = \beta(\lambda_\infty - \lambda_t)dt + \alpha dN_t, \tag{14}$$

where, $\mu^\star = \mathbb{E}\left[e^{J_x} - 1\right]$ and $\bar{\delta}^{\mathbb{Q}} = \bar{\delta} - \varphi_\delta/\gamma$.

We remark again that the jump process appearing in the log-returns dynamics is a compensated compound Poisson process and that the oil price (properly discounted) is now a local martingale with respect to $\mathbb{Q}$, in order to be consistent with the no-arbitrage principle. The model we are proposing belongs to the class of affine models, which means that we know the characteristic function in closed-form. This feature of affine models is very important in order to price financial derivatives as we shall see below. The payoff of a futures contract $F(t,\tau)$, with time to maturity $\tau = T - t$, is given by the usual relation:

$$F(t,\tau) = \mathbb{E}^{\mathbb{Q}}[S_T \mid \mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}}[e^{X_T} \mid \mathcal{F}_t], \tag{15}$$

where $X_T = \ln(S_T)$. We start by considering the moment generating function (MGF) of $X_T$:

$$G(w, X_t, V_t, \delta_t, \lambda_t, t, \tau) = \mathbb{E}\left[e^{wX_T} \mid \mathcal{F}_t\right]$$

Now, since under the risk-neutral measure the model structure is the same we drop the $\mathbb{Q}$ superscript to lighten the notation, and by Feynman-Kac theorem we obtain the following PDE:

$$\begin{aligned} &- G_\tau + \left( r - \frac{1}{2} V_t - \lambda_t \mu^\star - \delta_t \right) G_x + \frac{1}{2} V_t G_{xx} + k(\bar{V} - V_t) G_v + \\ &+ \frac{1}{2}\sigma_v^2 V_t G_{vv} + \rho_v \sigma_v V_t G_{xv} + \beta(\lambda_\infty - \lambda_t) G_\lambda + \gamma(\bar{\delta} - \delta_t) G_\delta + \frac{1}{2}\sigma_\delta^2 G_{\delta\delta} \\ &+ \lambda_t \int \left[ G(w, X_t + J_x, V_t + J_v, \delta_t, \lambda_t + \alpha, t, T) - G(w, X_t, V_t, \delta_t, \lambda_t, t, T) \right] \nu(dJ_x, dJ_v) = 0 \end{aligned} \tag{16}$$

with terminal condition $G_T = \exp(wX_T)$. Now, we guess a solution of the form:

$$G(w, X_t, V_t, \delta_t, \lambda_t, t, \tau) = \exp\left\{wX_t + A(w,\tau) + B(w,\tau)V_t + C(w,\tau)\delta_t + D(w,\tau)\lambda_t\right\} \tag{17}$$

subject to $A(0) = 0$, $B(0) = 0$, $C(0) = 0$ and $D(0) = 0$. From the PDE (16) we obtain the following system of ODEs:

$$\begin{cases} \frac{\partial A(w,\tau)}{\partial \tau} = rw + \beta\lambda_\infty D(w,\tau) + k\bar{V}B(w,\tau) + \gamma\bar{\delta}C(w,\tau) + \frac{1}{2}\sigma^2 C^2(\tau), \\ \frac{\partial B(w,\tau)}{\partial \tau} = -\frac{1}{2}(w - w^2) - (k - \rho\sigma_v w)B(w,\tau) + \frac{1}{2}\sigma_v^2 B^2(w,\tau), \\ \frac{\partial C(w,\tau)}{\partial \tau} = -w - \gamma C(w,\tau), \\ \frac{\partial D(w,\tau)}{\partial \tau} = -\beta D(w,\tau) + \int \left[e^{wJ_x + B(w,\tau)J_v + D(w,\tau)\alpha} - 1\right] \nu(dJ_x, dJ_v) - \mu^\star w. \end{cases} \tag{18}$$

Now, the futures price is given simply by the MGF computed in $w = 1$; then the solution of the second ODE is $B(\tau) = 0$, since it is a Riccati equation without the constant term. Due to this result we also have $D(\tau) = 0$. Hence, as we said before, by construction future prices are not affected by volatility and jumps. Therefore, the solution for the log-futures is of the form:

$$\ln F(t,\tau) = \ln S_t + A^{\mathbb{Q}}(\tau) + C^{\mathbb{Q}}(\tau)\delta_t, \tag{19}$$

where,

$$A^{\mathbb{Q}}(\tau) = r\tau + \frac{\delta^{\mathbb{Q}}\left(-\gamma\tau - e^{-\gamma\tau} + 1\right)}{\gamma} + \frac{\sigma_\delta\left(0.5\gamma\tau - 0.25e^{-2\gamma\tau} + e^{-\gamma\tau} - 0.75\right)}{\gamma^3}, \tag{20}$$

$$C^{\mathbb{Q}}(\tau) = \frac{e^{-\gamma\tau} - 1}{\gamma}. \tag{21}$$

## 2.3 An Optimal Dynamic Hedging Strategy

Hedging in crude oil market is an important issue not only for producers, but also for energy traders and investors (Billio *et al.*, 2018). Indeed, price fluctuations lead to an increase in volatility, and so the risk coming from investing in the spot market need to be mitigated. In particular, the natural way to hedge a long (short) position in the spot market is to sell (buy) a certain number of futures contracts. The quantity of futures needed to cover a spot position is called *hedge ratio*. The determination of the optimal hedge ratio depends on the chosen objective function. In the literature, the hedge ratio is modeled as a time-varying variable (see Kroner

and Sultan, 1993; Liu *et al.*, 2014; Billio *et al.*, 2018; Chang *et al.*, 2011; Batten *et al.*, 2019; Alizadeh *et al.*, 2008), which minimizes the variance of the portfolio $\Pi = S - hF$, obtaining as solution

$$h^{\star}_{t-1} = \frac{\mathrm{Cov}_t(S,F)}{\mathrm{Var}_t(F)}. \tag{22}$$

At this point, it is important to remark that our approach is different from that usually considered in the literature. In particular, the most popular choice is given by discrete time models like multivariate GARCH models (Chang *et al.*, 2011; Batten *et al.*, 2019) and Markov switching models (Billio *et al.*, 2018; Alizadeh *et al.*, 2008). On the other hand, continuous time models have received much less attention and one example of such an approach is given by Liu *et al.* (2014). In their paper, related to industrial metals' market, they specify one dynamics for the spot price and one for the futures and then the estimation is carried out separately. However, our model provides a direct link between spot and futures, meaning that in our estimation framework we cannot avoid no arbitrage issues and we need to resort to a risk-neutral argument. In view of possible extensions of the present estimation method including different derivatives contracts, European options for example, the risk-neutral approach represents the most convenient and natural modelling framework. In the present setting, the risk-neutral approach provides the necessary consistency relations between spot and futures prices.

Now, let us properly specify the hedging portfolio. In particular, the log-spot price is given by:

$$dX_t = \left(r - \frac{1}{2}V_t - \lambda_t\mu^{\star} - \delta_t\right)dt + \sqrt{V_t}dW_{x,t} + d\widetilde{J}_{x,t}.$$

Moreover the latent states are described by the following equations:

$$\begin{aligned}
dV_t &= k(\bar{V} - V_t)dt + \sigma_v\sqrt{V_t}\left(\rho_v dW_{x,t} + \sqrt{1-\rho_v^2}dW_{v,t}\right) + dJ_{v,t},\\
d\delta_t &= \gamma(\bar{\delta} - \delta_t)dt + \sigma_\delta dW_{\delta,t},\\
d\lambda_t &= \beta(\lambda_\infty - \lambda_t)dt + \alpha dN_t.
\end{aligned}$$

Then, By applying Itô's lemma on the log-futures pricing function $f(X,V,\lambda,\delta,t,T) = X_t +$

$A(\tau) + C(\tau)\delta_t$ we get the dynamics of $Y_t$:

$$dY_t = \Big[r - \frac{1}{2}V_t - \lambda_t\mu^\star - \delta_t - A'(\tau) - C'(\tau)\delta_t + C(\tau)\gamma(\bar{\delta} - \delta_t)\Big]dt$$
$$+ \sqrt{V_t}dW_{x,t} + C(\tau)\sigma_\delta dW_{\delta,t} + d\widetilde{J}_{x,t}$$

Then, the MV hedging ratio is given by

$$h^\star = \frac{V_t + (\mu_J^2 + \sigma_J^2)\lambda_t}{V_t + C^2(\tau)\sigma_\delta^2 + (\mu_J^2 + \sigma_J^2)\lambda_t}. \tag{23}$$

Additional details are given in subsection 5.1 and in Appendix B, where we also introduce a skewness-type objective function.

# 3 Parameters Estimation Method

In order to estimate the parameters using real data, we cast our model in a state-space form and apply a simple Euler scheme. Denoting by $\Delta t$ a small time interval, the observation equation for stock prices is given by

$$\ln S_t = \ln S_{t-1} + \left(\mu - \frac{1}{2}V_{t-1} - \lambda_{t-1}\mu^\star - \delta_{t-1}\right)\Delta t + \sqrt{V_{t-1}\Delta t}\, W_t + J_x \Delta N_t, \tag{24}$$

where, $J_x \sim \mathcal{N}(\mu_J, \sigma_J^2)$, $W_t \sim \mathcal{N}(0,1)$ and $\Delta N_t = N_t - N_{t-1} \sim \text{Bernoulli}(\lambda_{t-1}\Delta t)$.

In this work we also consider futures prices with $n$ maturities. As usual in the literature, we observe derivatives data with measurement errors (Eraker, 2004, Fulop and Li, 2019, etc.); then we have an additional observation equation:

$$\ln F(t,T)^O = \ln F(t,T)^M + \epsilon_t,$$

where, $\ln F(t,T)^O$ is a vector of observed futures prices at time $t$ with maturity $T$, and $\ln F(t,T)^M$ are corresponding prices obtained with Equation (19). The measurement errors follow a multivariate normal, i.e. $\epsilon_t \sim \mathcal{N}(0,\Omega)$, with $\Omega = \Sigma\Sigma^T$ and $\Sigma = \sigma_e I_n$.

Our model features latent states, which cannot be observed. The discretized version of Equations (6)-(7)-(8) reads as follows:

$$V_t = V_{t-1} + k(\bar{V} - V_{t-1})\Delta t + \sigma_v \sqrt{\Delta t V_{t-1}} Z_t + J_v \Delta N_t \tag{25}$$

$$\delta_t = \delta_{t-1} + \gamma(\bar{\delta} - \delta_{t-1})\Delta t + \sigma_\delta \sqrt{\Delta t} W_{\delta,t} \tag{26}$$

$$\lambda_t = \lambda_{t-1} + \beta(\lambda_\infty - \lambda_{t-1})\Delta t + \alpha \Delta N_t \tag{27}$$

where, $Z_t$ is a standard normal correlated with $W_t$ in Equation (24), $J_v \sim \text{Exp}(\mu_v)$ and $W_{\delta,t}$ is an independent standard normal. To address the estimation of the model we rely on Sequential Monte Carlo Methods (SMC). In particular, we shall consider the two-stage density tempered SMC, recently proposed by Fulop and Li (2019). This method provides a direct route from the prior to the posterior using all data in a row. To be more precise, by starting from the prior, we temper the distribution in order to gradually reach the final posterior through the tempering coefficient $\xi$, which is chosen adaptively to ensure satisfactory particle diversity, as in Fulop and Li (2019). In this framework we proceed in two steps: in the first stage we use a small number of state particles $M_1$ in order to provide a coarse exploration of the posterior and then in the second stage we increase the number of particles $M_2$ to correct the error in the likelihood estimation. Usually, the second stage is much faster then the first one leading to an overall advantage with respect to an SMC sampler which runs throughout with a fixed $M_2$. The other main feature of this procedure is the smoothing of the likelihood in the particle filter and the introduction of common random numbers in order to reduce the variance in the likelihood estimation, which is crucial to obtain reliable estimates. Furthermore, our estimation approach delivers an approximation of the marginal likelihood,

$$p(y_{1:T}) = \int p(y_{1:T} \mid \Theta) p(\Theta) d\Theta,$$

which can be used to construct Bayes factors for model comparison. More precisely, for two models $M_1$ and $M_2$ the Bayes factor is given by the ratio of their marginal likelihoods,

$$BF_{1,2} = \frac{p(y_{1:T} \mid M_1)}{p(y_{1:T} \mid M_2)}.$$

The Bayes factor does not rely on asymptotic distribution theory and provides a simple way to evaluating different models. Besides, it contains a penalty for using more parameters. For

a complete exposition about the two-stage SMC sampler algorithm we refer to Fulop and Li (2019).

# 4 Empirical Application

## 4.1 Data Description

In this section we present the data we are going to examine in our empirical analysis. In particular, we consider WTI Cushing (Oklahoma) crude oil spot and futures quotations obtained from Bloomberg[2]. In Figure 2 we plot a time series of daily log-returns: $Y_t = \log(S_t/S_{t-1})$, ranging from 08/01/2008 to 31/12/2018, thus 2767 spot data are considered. The data cover some crucial historical periods as the global financial crisis, when oil prices experienced huge fluctuations between 2008 and 2010. Moreover, in the middle of 2014, price started declining due to a significant increase in oil production in USA, and declining demand in the emerging countries. Besides, from 2016 to 2018, complex negotiations with OPEC led to higher variability in oil prices.

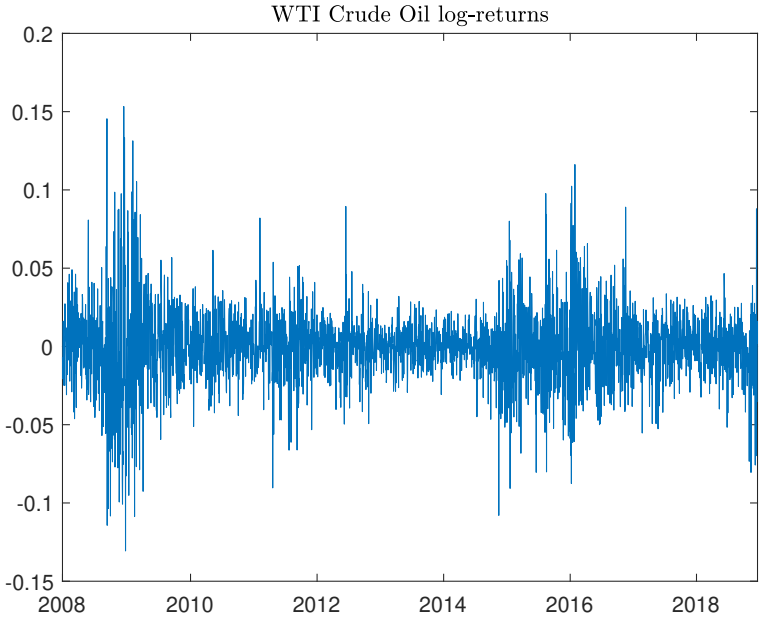


Figure 2: WTI crude oil log-returns from 08/01/2008 to 31/12/2018.

[2]spot data ticker: `USCRWTIC`, futures data ticker: `CLm`, for $m = \{3, 6, 9, 12, 18, 21, 24\}$.

First, a preliminary statistical analysis is performed on log-returns. To this end in Table 1 we show some descriptive statistics and the result of a Jarque-Bera (JB) normality test. In addition a QQ-plot is provided in Figure 3. This simple analysis suggests that observations do not come from a Gaussian distribution.

| **Statistics** | **Log-return** |
|---|---|
| Observations | 2767 |
| Mean | -2.7179e-04 |
| Standard Deviation | 0.0245 |
| Skewness | 0.1181 |
| Kurtosis | 7.3057 |
| Min value | -0.1307 |
| Max value | 0.1533 |
| JB test | Rejected |

Table 1: Descriptive statistics and JB test on WTI crude oil log returns observed with daily frequency over the period 08/01/2008 to 31/12/2018.

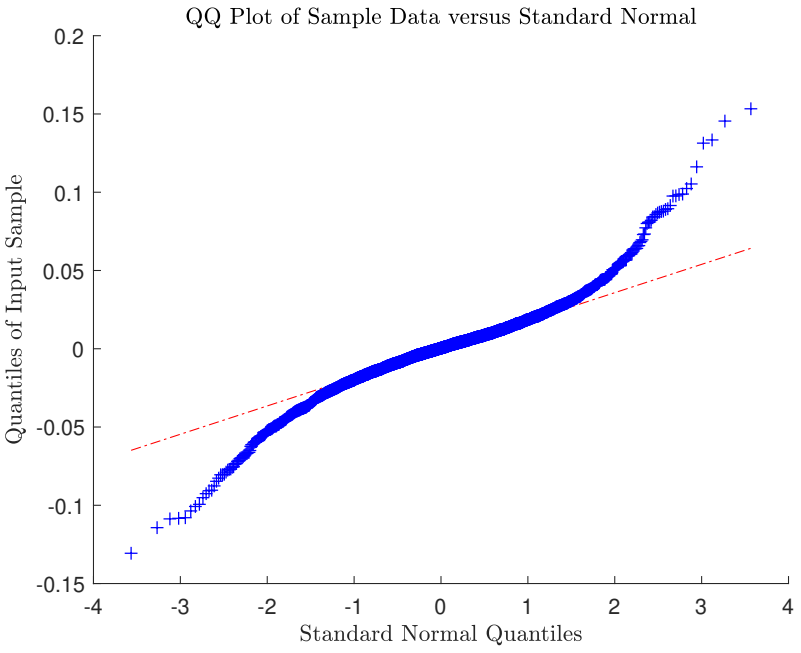


Figure 3: Q-Q Plot for daily frequency returns on WTI crude oil.

On the other hand, we consider futures contracts written on WTI crude oil. In Figure 4 we show log-futures prices from 08/01/2008 to 31/12/2018, for maturities $\tau$ equal to 3, 12 and 24 months. For estimation purpose, we retain 8 futures contracts with fixed maturities ranging from 3 to 24 months. Contracts with more than 2 years of maturity are less liquid (Lai and Mellios, 2016), therefore they are not considered in our estimation procedure. As for spot observations, we conduct a statistical analysis on the whole futures data in Table 2.

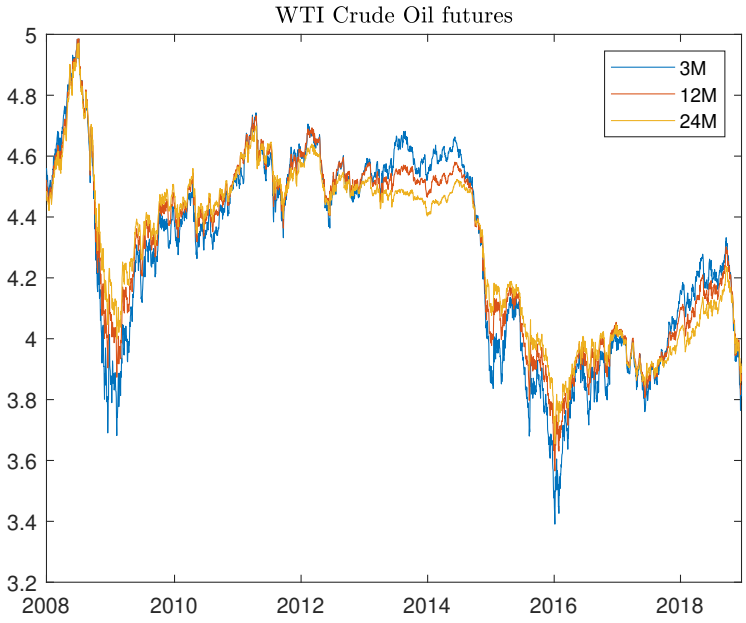


Figure 4: WTI futures contracts from 08/01/2008 to 31/12/2018.

| **Statistics** | **CL3** | **CL6** | **CL9** | **CL12** | **CL15** | **CL18** | **CL21** | **CL24** |
|---|---|---|---|---|---|---|---|---|
| Mean | 4.2842 | 4.2968 | 4.3019 | 4.3041 | 4.3047 | 4.3046 | 4.3043 | 4.3039 |
| Standard Dev | 0.3217 | 0.3058 | 0.2956 | 0.2879 | 0.2814 | 0.2760 | 0.2713 | 0.2673 |
| Skewness | -0.3142 | -0.2751 | -0.2547 | -0.2422 | -0.2337 | -0.2277 | -0.2224 | -0.2201 |
| Kurtosis | 2.0059 | 1.9591 | 1.9465 | 1.9358 | 1.9328 | 1.9388 | 1.9475 | 1.9598 |
| Min value | 3.3908 | 3.4825 | 3.5293 | 3.5656 | 3.5943 | 3.6168 | 3.6368 | 3.6548 |
| Max value | 4.9845 | 4.9894 | 4.9895 | 4.9858 | 4.9822 | 4.9782 | 4.9747 | 4.9716 |

Table 2: Descriptive statistics on WTI crude oil log futures prices observed with daily frequency over the period 08/01/2008 to 31/12/2018.

## 4.2 Estimation Results

In this subsection we present the results obtained from the estimation procedure. In our empirical application, for the two-stage SMC sampler, we set the number of parameter particles $N$ equal to 1000, the number of state particles at first stage $M_1$ equal to 30, and the number of state particles $M_2$ at second stage equal to 500. The algorithm is initialized using the priors in Table 3 and with $\Sigma = \sigma_e I_n$. The choice of the hyper-parameters of the prior distributions is based on calibration using the training sample approach, which is widely used to calibrate the objective priors (Fulop and Li, 2013, 2019; Fulop *et al.*, 2015). Notably, we find that most parameters are not sensitive to the selection of the priors. Moreover, we take the random walk proposal, trigger the resample-move step when the ESS reaches $N/2$, and then keep moving until the cumulative average acceptance rate across the population reaches 2.

| $\Theta$ | Dist | Support | $(\mu_0, \sigma_0)$ | $\Theta$ | Dist | Support | $(\mu_0, \sigma_0)$ |
|---|---|---|---|---|---|---|---|
| $\mu$ | Normal | $(-\infty, \infty)$ | (0.02, 0.15) | $\varphi_\delta$ | Tr. Normal | $(-\infty, \infty)$ | (0.5, 0.1) |
| $\mu_J$ | Tr. Normal | $(0, \infty)$ | (-0.02, 0.08) | $\gamma$ | Normal | $(0, \infty)$ | (0.9, 0.5) |
| $\sigma_J$ | Tr. Normal | $(0, \infty)$ | (0.05, 0.1) | $\bar{\delta}$ | Tr. Normal | $(0, \infty)$ | (0.15, 0.05) |
| $k$ | Tr. Normal | $(0, \infty)$ | (3.0, 4.0) | $\sigma_\delta$ | Tr. Normal | $(0, \infty)$ | (0.25, 0.1) |
| $\bar{V}$ | Tr. Normal | $(0, \infty)$ | (0.05, 0.06) | $\alpha$ | Tr. Normal | $(0, \infty)$ | (2.5, 1.5) |
| $\sigma_v$ | Tr. Normal | $(0, \infty)$ | (0.25, 0.25) | $\beta$ | Tr. Normal | $(0, \infty)$ | (4.0, 3.0) |
| $\rho_v$ | Tr. Normal | $[-1, 1]$ | (-0.7, 0.5) | $\lambda_\infty$ | Tr. Normal | $(0, \infty)$ | (2, 0.8) |
| $\mu_v$ | Tr. Normal | $(0, \infty)$ | (0.02, 0.1) | | | | |

Table 3: Priors specification.

To appreciate the efficiency of our estimation method, we can have a look at the acceptance rates related to the moving step in Figure 5. The star-line indicates the acceptance rates from the first stage and the circle line those from the second stage. We clearly see that the acceptance rates remain high during both the first and second stage. Furthermore, as demonstrated in Fulop and Li (2019), the number of density-bridging iterations is much smaller in the second stage with respect to the first stage.

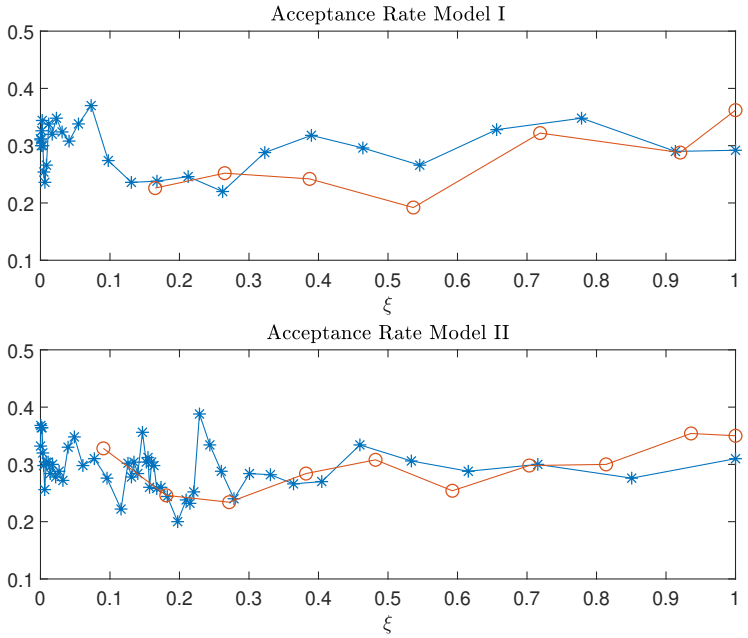


Figure 5: The figure plots the last acceptance rate in moving steps at each density-bridging iteration with respect to $\xi_i$ for model I (upper) and model II (bottom). In the algorithm, $\xi$ is automatically selected using a grid search approach. The blue line refers to the first stage, while the orange line refers to the second stage.

Figures 6 and 7 plot the filtered volatility, convenience yield and jump intensity obtained by running our smooth particle filter (see Appendix A) at the posterior mean. In particular, volatility is quite persistent and in periods when prices fall down we observe a rise in volatility according to the well known leverage effect. The convenience yield is moving in the same direction of oil prices, which is consistent with its definition and provides a clear economic intuition. A large amount of literature devoted to commodities provides a confirmation and an explanation of this behavior, we just mention Alquist *et al.* (2014) and Lautier (2009) among many other contributions on this topics, some based on Normal Backwardation Theory (Litzenberger and Rabinowitz, 1995), some on Theory of Storage (Casassus *et al.*, 2005). The intensity process present a self-exciting behaviour; this will be confirmed later when focusing on the parameters estimates.

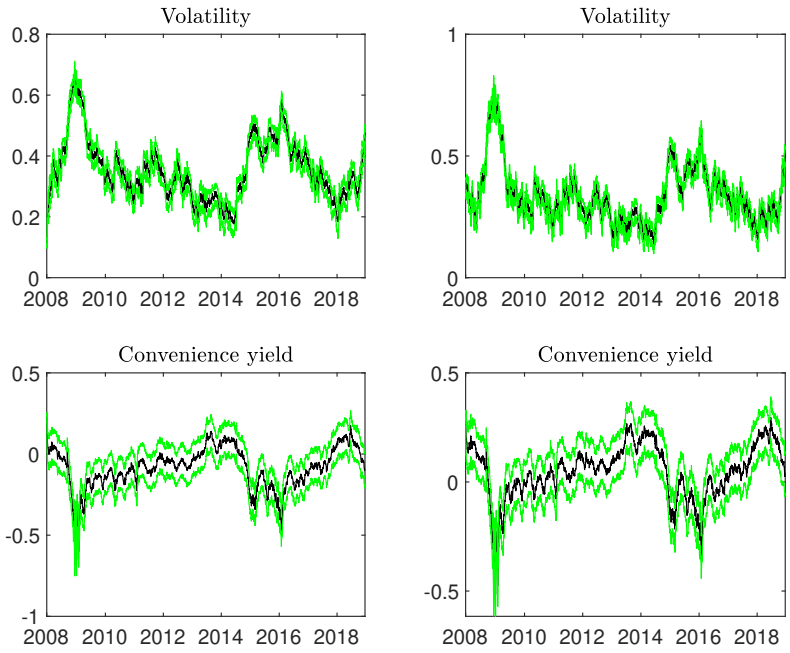


Figure 6: The figure presents the filtered volatility and convenience yield from model I (left) and II (right). The posterior mean and $(5, 95)\%$ quantiles are reported at each time point.

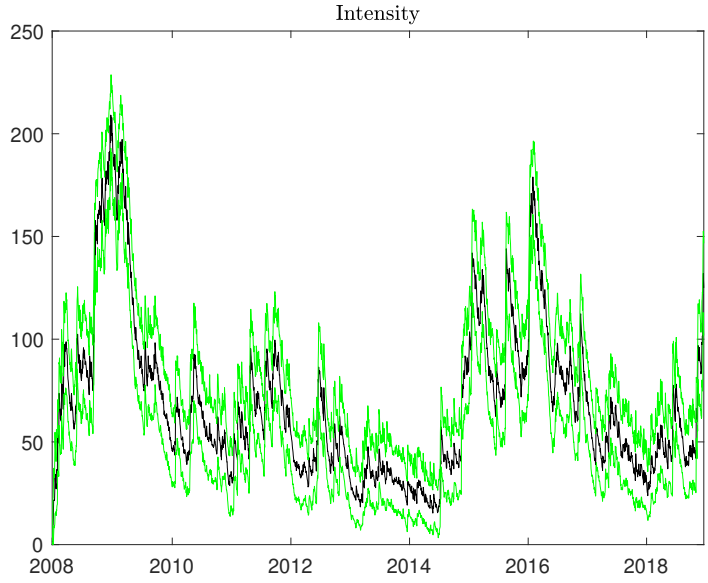


Figure 7: The figure presents the filtered jump intensity from model I. The posterior mean and $(5, 95)\%$ quantiles are reported at each time point.

Table 4 presents the parameter estimates for the two models obtained with the two-stage SMC sampler. As a first remark, when self-excitation is considered, volatility is less persistent. Indeed, $k$ is equal to 9.2692 (2.1103) in Model I and 2.9157 (0.4418) in Model II. One possible explanation for this is that much of the variation is due to the self-exciting jump intensity, dampening the contribution of volatility. However, the vol-of-vol parameter $\sigma_v$ is equal to 0.3566 (0.0960) in Model I and 0.8498 (0.0656) in Model II, confirming that volatility dynamics in oil market is clearly stochastic. Second, in line with the previous literature (Larsson and Nossman, 2011; Brooks and Prokopczuk, 2013), we find evidence of volatility jumps. In particular, the estimate for $\mu_v$ is 0.0223 (0.0063) in Model I and 0.0311 (0.0089) in Model II. Third, the posterior mean of the parameter controlling the self-exciting effect, $\alpha$, is equal to 23.4601 (4.1515); then $\alpha$ is well identified and constitutes a key feature of the jump dynamics in the oil market. Fourth, the convenience yield dynamics is pretty much the same within the two models. The parameters are well identified and the convenience yield risk premium is statistically significant, which confirms previous studies in the literature.

Furthermore, it is possible to compare the models by looking at the marginal likelihood and the log Bayes factor. In both cases we can say that Model I performs better than Model II. For instance, the log Bayes factor[3] of Model I with respect to Model II is 9.1271, which means that Model I is decisively better than Model II in fitting the data.

[3]For any two given models, $M_1$ and $M_2$, if the value of the log Bayes factor is between 0 and 1.1, $M_1$ is barely worth mentioning; if it is between 1.1 and 2.3, $M_1$ is substantially better than $M_2$; if it is between 2.3 and 3.4, $M_1$ is strongly better than $M_2$; if it is between 3.4 and 4.6, $M_1$ is very strongly better than $M_2$; and if it is larger than 4.6, $M_1$ is decisively better than $M_2$.

| | Model I | | Model II | |
|---|---|---|---|---|
| $\Theta$ | Mean | Std | Mean | Std |
| $\mu$ | 0.1451 | 0.0944 | 0.0109 | 0.0691 |
| $\mu_J$ | 0.0022 | 0.0047 | 0.0118 | 0.0827 |
| $\sigma_J$ | 0.0408 | 0.0038 | 0.1104 | 0.0537 |
| $k$ | 9.2692 | 2.1103 | 2.9157 | 0.4418 |
| $\bar{V}$ | 0.0131 | 0.0067 | 0.1309 | 0.0134 |
| $\sigma_v$ | 0.3566 | 0.0960 | 0.8498 | 0.0656 |
| $\rho_v$ | -0.7628 | 0.1645 | -0.4621 | 0.0414 |
| $\gamma$ | 0.6708 | 0.0041 | 0.6721 | 0.6721 |
| $\bar{\delta}$ | 0.2304 | 0.0086 | 0.2639 | 0.0079 |
| $\sigma_\delta$ | 1.7690 | 0.0183 | 1.7646 | 0.0197 |
| $\beta$ | 30.0566 | 3.7743 | (—) | (—) |
| $\alpha$ | 23.4601 | 4.1515 | (—) | (—) |
| $\lambda_\infty$ | 1.3921 | 0.0079 | 8.0900 | 0.0518 |
| $\varphi_\delta$ | 0.1272 | 0.0052 | 0.1500 | 0.0048 |
| $\mu_v$ | 0.0223 | 0.0063 | 0.0311 | 0.0089 |
| MLLH | 4.1901e+4 | | 4.1550e+4 | |

Table 4: Parameter estimates for model I and II. For each parameter, the posterior mean and standard deviation are reported. The last row reports the marginal log likelihood estimated with the smooth particle filter.

Now, we can provide some additional insights regarding the convenience yield dynamics and the forecasting ability of the models in terms of futures pricing. In particular, with this kind of models it is common to observe a lack of fit in the filtered convenience yield process (Carmona and Ludkovski, 2004). To check if our approach could give a reliable estimate of this variable, we follow Carmona and Ludkovski (2004) and compute the implied convenience yield using the estimated parameters together with the price of traded futures contracts $F(t, T_i)$. Then, we compare the filtered convenience yield from our smooth particle filter with the implied one. The result of this exercise is shown in Figure 8 and confirms that our filtering method is able to well reproduce the convenience yield dynamics implied from data (especially for longer maturities).

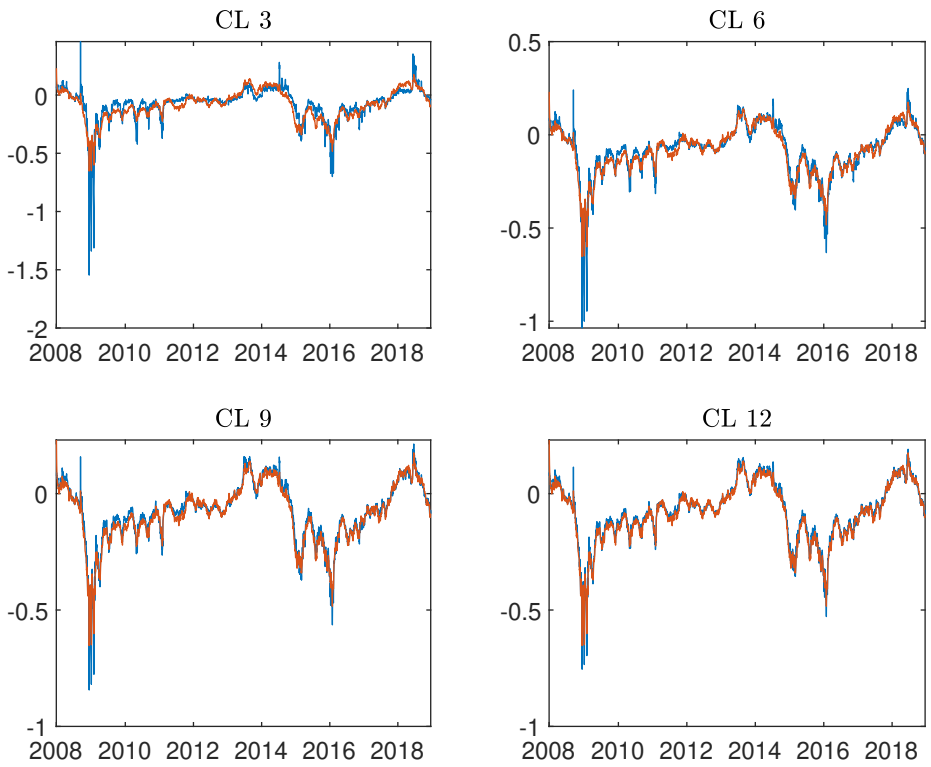


Figure 8: The figure compares the filtered convenience yield (orange) against the implied convenience yield (blue) computed using futures contracts with $\tau = \{3, 6, 9, 12\}$ and the posterior mean parameter estimates from Model I.

Second, we check the pricing errors of futures contracts when Model I and Model II are considered. For example, in Figure 9, we can see how the models behave in terms of futures pricing out of the sample, i.e. from 02/01/2019 to 17/01/2020 (263 observations). To this end we run our smooth PF on the new data using the parameters obtained during the previous estimation and then we filter out the convenience yield. As we can see the results are good with both models, with better performances for Model I.

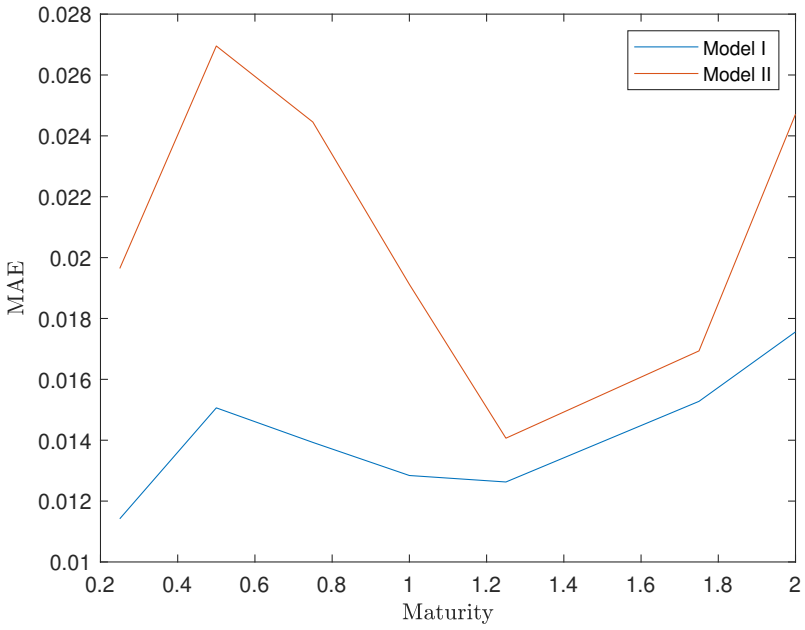


Figure 9: The figure presents the out of sample Mean Absolute Error (MAE) in terms of pricing futures contracts with Model I and Model II.

# 5 Financial Application: Dynamic Hedging

In our application we follow previous studies (e.g. Liu *et al.*, 2014; Billio *et al.*, 2018) and consider as hedging instrument the most liquid futures, i.e. the futures with the shortest maturity, which in our data set are the CL3 contracts. Thus, we compute a daily hedge ratio according to the MV approach. In Figure 10 the optimal time varying hedge ratio $h_t^\star$ computed with Model I and Model II is shown. The time varying hedge ratio indicates that the portfolio should be frequently re-balanced as the market conditions change. In particular, during the global financial crisis (2008-2010) we observe a positive jump in the hedge ratio, which means that investors are more cautious and they prefer to hedge more. Then, from 2011 to 2014 the market is less volatile and the hedge ratios are smaller providing to the hedgers a lower exposition on futures. Finally, from 2015-2017 we observe again an increase in the optimal hedging strategy due to another period of market turmoil. Then, we provide some descriptive statistics about the hedge ratios in Table 5.

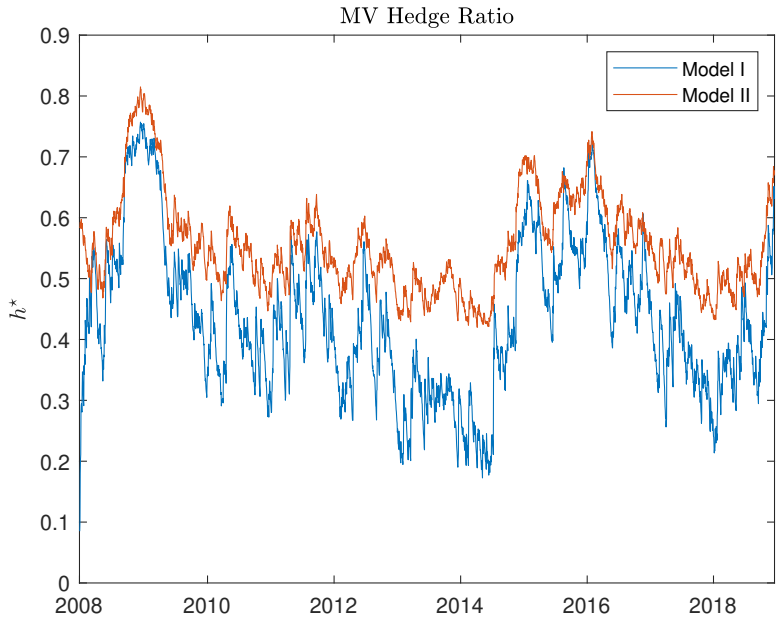


Figure 10: The figure presents the optimal MV hedge ratio time series for Model I (blue) and Model II (orange).

| | Mean | Std Dev | Skewness | Kurtosis | Min | Max |
|---|---|---|---|---|---|---|
| MV Hedge I | 0.4326 | 0.1250 | 0.4134 | 2.7714 | 0.0851 | 0.7571 |
| MV Hedge II | 0.5552 | 0.0801 | 0.9150 | 3.6083 | 0.4199 | 0.8149 |

Table 5: MV Hedge ratio statistics.

In the literature it is common to evaluate a particular hedging strategy in terms of variance reduction and utility improvements with respect to the un-hedged position (Kroner and Sultan, 1993; Alizadeh *et al.*, 2008; Batten *et al.*, 2019). Therefore, we consider the following measure of hedging effectiveness:

$$HE_1 = 1 - \left[\frac{\text{Var}(\Pi_h)}{\text{Var}(\Pi_{un})}\right], \tag{28}$$

where the hedging portfolios $\Pi_h$ are computed by using percentage log-returns of spot and futures and the corresponding optimal hedge ratios. Hence, we are evaluating the variance reduction with respect to the un-hedged portfolio $\Pi_{un}$, which is composed only by the spot position. If $HE_1 = 0$ we do not reduce risk at all, whereas $HE_1 = 1$ imply a 100% reduction in

the variance. According to our previous findings about the hedge ratio variability, we consider the whole sample and two different sub-periods: 2008-2010 and 2015-2017. In this way we are able to assess the hedging performances related to specific turbulent periods. The results are shown in Table 6.

| Interval | Model I | Model II |
|---|---|---|
| 2008-2018 | 0.6030 | 0.5554 |
| 2008-2010 | 0.6394 | 0.6117 |
| 2015-2017 | 0.6568 | 0.6040 |

Table 6: The table presents the variance reduction with respect to the un-hedged portfolio.

Overall, we obtain clear improvements with respect to the un-hedged position. For instance, during the whole sample, we reduce the portfolio variance by 60.30% with Model I and by 55.54% with model II. Moreover, even during periods of financial turmoil we are able to reduce significantly the variance of the portfolio. This confirms the importance of hedging for producers and investors.

According to Alizadeh *et al.* (2008), hedging effectiveness is more appropriately assessed by considering the economic benefits from hedging using the hedger's utility function. Then, if $\xi$ represents the risk aversion of an investor and $R_h$ is the return on the hedged portfolio, the expected utility function is given by

$$\mathbb{E}[U(R_h)] = \mathbb{E}[R_h] - \xi \text{Var}[R_h]. \tag{29}$$

By assuming that expected returns from the hedged portfolio are equal to zero and the degree of risk aversion is 4[4], we compute the realized utility for Model I, Model II and the un-hedged portfolio in Table 7.

| Interval | Model I | Model II | Un-hedged |
|---|---|---|---|
| 2008-2018 | -9.5137 | -10.6541 | -23.9646 |
| 2008-2010 | -19.3052 | -20.7915 | -53.5398 |
| 2015-2017 | -9.3984 | -10.8453 | -27.3873 |

Table 7: The table presents the realized MV utility for Model I, Model II and the un-hedged portfolio.

[4]These assumptions are in line with most empirical studies as Kroner and Sultan (1993); Alizadeh *et al.* (2008).

Now, some comments are needed. First, by adopting the optimal hedging strategy we obtain substantial utility improvements with respect to the un-hedged position, no matter which sub-period is considered. Second, during the global financial crisis we get the worst results and this is due to the increased portfolio variance. Third, Model I is performing better than Model II, both regarding variance reduction and utility improvement, by confirming the importance of a more elaborate jump structure with respect to the standard Poisson framework with constant intensity.

In the next subsection we are going to explore if the inclusion of high order effects in the objective function could provide some improvements in the present hedging application.

## 5.1 Higher Order Hedging

The evidence of jumps and stochastic volatility in the oil market raises a question about the adequacy of a minimum variance objective function. Indeed, by adopting the MV approach we are not fully exploiting the distributional properties of these models, i.e. we are neglecting higher order effects, which could influence the hedging ratio. For instance, the relevance of skewness in characterizing risk preferences has been pointed out in a rich amount of literature: we mention the papers by Post *et al.* (2008), Chiu (2010), Dahlquist *et al.* (2017) and Kraus and Litzenberger (1976). Therefore, we consider the minimization of the following objective function:

$$\min_h [\mathrm{Var}_t(\Pi_t) - \eta \mathrm{Asy}_t(\Pi_t)], \tag{30}$$

where, $\Pi_t = X_t - hY_t$ is the portfolio formed by $X_t = \log(S_t)$ and $Y_t = \log(F_{t,\tau})$; $\eta$ is a constant risk-aversion parameter and $\mathrm{Asy}[x] = \mathbb{E}[(x - \mathbb{E}[x])^3]$.

In order to compute the optimal hedging ratio we need $\mathrm{Var}_t[\Pi_t]$ and $\mathrm{Asy}_t[\Pi_t]$. Hence, in analogy with Liu *et al.* (2014) we can proceed by computing the instantaneous conditional moments. The detailed derivation of the hedging ratio is given in Appendix B. Once we have obtained the expression for the hedging portfolio we take the first derivative with respect to $h$ and set it equal

to zero. In the end we get two solutions for $h^\star$:

$$h^\star = \pm \frac{1}{2\left(3\eta\lambda_t\mu_J^3 + 9\eta\lambda_t\mu_J\sigma_J^2\right)} \Big( \left(2C^2\sigma_\delta^2 - 6\eta\lambda_t\mu_J^3 - 18\eta\lambda\mu_J\sigma_J^2 + 2\lambda_t\mu_J^2 + 2\lambda_t\sigma_J^2 + 2V\right)^2 - 4\left(3\eta\lambda_t\mu_J^3 + 9\eta\lambda_t\mu_J\sigma_J^2\right)\left(3\eta\lambda_t\mu_J^3 + 9\eta\lambda_t\mu_J\sigma^2 - 2\lambda_t\mu_J^2 - 2\lambda_t\sigma_J^2 - 2V_t\right) - 2C^2\sigma_\delta^2 + 6\eta\lambda_t\mu_J^3 + 18\eta\lambda_t\mu_J\sigma_J^2 - 2\lambda_t\mu_J^2 - 2\lambda_t\sigma_J^2 - 2V_t\Big)^{1/2}. \tag{31}$$

Since crude oil spot and futures are positively correlated, from now on, we shall consider only the positive solution of (31). As in the previous section, we compute daily hedge ratios according to our optimal strategy. Nevertheless, in Figure 11 we observe a surprising result: the MV and the Asy-Variance hedging strategies give substantially the same hedge ratio. This mean that the contribution given by the third-order moment is negligible in computing the optimal hedging strategy.

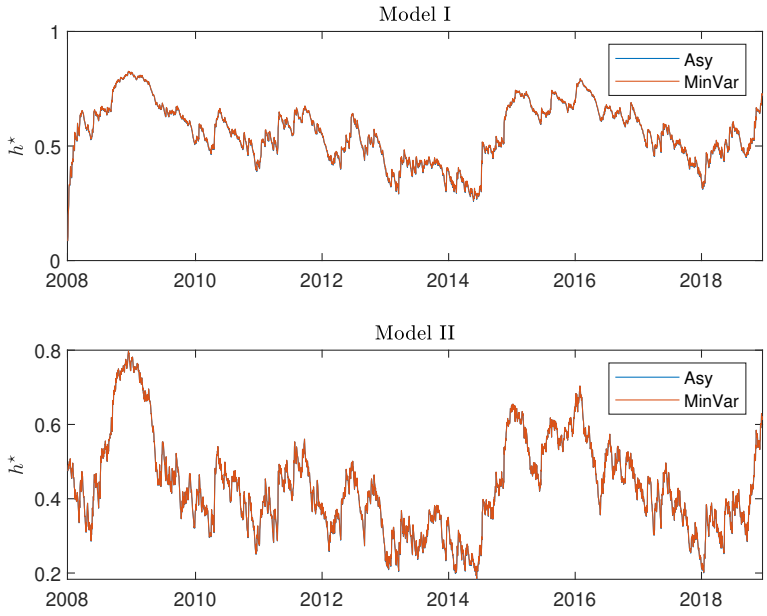


Figure 11: The figure presents the optimal MV and Asy-Variance hedge ratios time series for Model I (top) and Model II (bottom).

Given this result, we include the skewness instead of the third order central moment; then

the optimization problem reads as follows

$$\min_{h}[\mathrm{Var}_t(\Pi_t) - \eta \mathrm{Skew}_t(\Pi_t)], \tag{32}$$

where,

$$\mathrm{Skew}(x) = \frac{\mathbb{E}[(x - \mathbb{E}[x])^3]}{\mathrm{Var}(x)^{3/2}}.$$

In this way we are not able to get a closed-form solution, so a numerical minimization method is adopted. Hence, the hedge ratio obtained numerically is compared to the MV hedging ratio in Figure 12, from which it is clear that the contribution of skewness is not negligible. This intuition will be formally tested empirically in what follows, where we denote the Variance-Skewness hedging by VS.

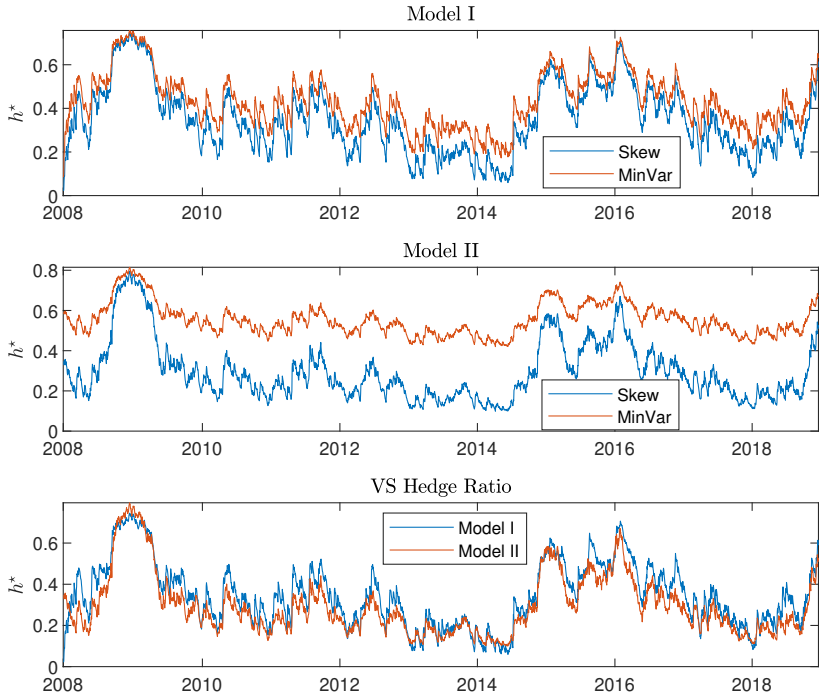


Figure 12: The figure presents the optimal MV and VS hedge ratios time series for Model I in the top panel and Model II in the middle panel. The bottom panel plots the VS hedge ratios for Model I (blue) and Model II (orange).

Now, in order to investigate the performance of this new strategy we need to consider a

metric which is coherent within the optimizations goals. Then, we resort again to the realized utility:

$$\mathbb{E}[U(R_h)] = -\text{Var}[R_h] + \eta \text{Skew}[R_h], \tag{33}$$

where, $\eta$ is the risk aversion parameter related to skewness and it is set equal to 4. The result of this computation is given in Table 8.

| Interval | Model I | Model II | Un-hedged |
|---|---|---|---|
| 2008-2018 | -1.0730 | -1.5485 | -5.5189 |
| 2008-2010 | -1.9980 | -2.5387 | -12.6846 |
| 2015-2017 | -1.9193 | -2.2930 | -5.7923 |

Table 8: The table presents the realized VS utility for Model I, Model II and the un-hedged portfolio.

From the results we infer that the realized utility is giving good performances. In particular, Model I is performing better than Model II in every scenario. Moreover, both models implied utilities achieve better results with respect to the un-hedged position. The explanation stems from the fact that positive skewness is increasing the utility of the investor.

In order to further investigate this phenomenon we test the MV and VS hedging strategies out of sample, i.e. we consider 263 observations from 02/01/2019 to 17/01/2020. Hence, we filter out the state variables by running our smooth particle filter by using the parameters estimated in the sample. Then, we compute the realized utility for each strategy. The results are shown in Table 9. By this way we can appreciate the performance of each strategy with respect to the un-hedged position, by matching their respective optimization goals.

| Utility | Model I | Model II | Un-hedhged |
|---|---|---|---|
| MV | -6.0292 | -4.0241 | -17.9647 |
| VS | -0.9270 | -1.1730 | -3.9147 |

Table 9: The table presents the out of sample MV and VS realized utility for Model I and Model II. The realized utility of the un-hedged portfolio is also reported.

The results indicate that both strategies give better results with respect to the un-hedged portfolio. In particular, we observe that Model I performs better than Model II when the VS utility is considered.

By resuming the previous results, it seems that the inclusion of high order effects lead to substantial improvements in the present hedging application with respect to the usual MV approach. Moreover, the model with self-exciting jump intensity outperforms the one with constant jump intensity.

# 6 Concluding Remarks

In this paper we propose a jump-diffusion model for oil price dynamics including self-exciting effects. Our model includes a stochastic dynamics for both the volatility and the yield coefficient. We provide an estimation method of Bayesian type based on a suitable adaptation of a SMC method proposed by Fulop and Li (2019). Our results show that the introduction of a jump intensity of Hawkes type improves the forecasting ability of this model with respect to a model without self-exciting effects. Moreover its affine feature allows to compute explicitly futures contracts prices providing a setting for accurate parameters estimation. Finally we compute an optimal hedging strategy based on futures trading. This optimal hedging strategy is obtained not only by minimizing variance, but by taking into account higher order moments, i.e. skewness. This optimal hedging strategy exhibits some interesting features and gives to the hedger better results with respect to the most popular MV approach. Furthermore, from the empirical application we infer that Model I, equipped with the self-exciting component, outperforms Model II in terms of hedging effectiveness and realized utility.

# Acknowledgements

The authors thank A. Fulop, R. Tédongap and B. Feunou for useful comments and suggestions. We also thank an anonymous referee, who provided remarks that helped ourselves to improve the quality of our original manuscript.

# References

Ait-Sahalia, Y., Chaco-Diaz, J. and Laeven, R. J. A. (2015) Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, **117**, 585–606.

Alizadeh, A. H., Nomikos, N. K. and Pouliasis, P. K. (2008) A Markov regime switching approach for hedging energy commodities. *Journal of Banking and Finance*, **32**, 1970—-1983.

Alquist, R., Bauer, H. and Diez de los Rios, A. (2014) What does the convenience yield curve tell us about the crude oil market? *Working Paper-Bank of Canada*, **2014-42**.

Batten, J. A., Kinateder, H., Szilagyi, P. G. and F., W. N. (2019) Hedging stocks with oil. *Energy Economics (forthcoming)*.

Baumeister, C. and Kilian, L. (2015) Forecasting the real price of oil in a changing world: A forecast combination approach. *Journal of Business and Economic Statistics*, **33**, 338–351.

Billio, M., Casarin, R. and Osuntuyi, A. (2018) Markov switching GARCH models for bayesian hedging on energy futures markets. *Energy Economics*, **70**, 545–562.

Brooks, C. and Prokopczuk, M. (2013) The dynamics of commodity prices. *Quantitative Finance*, **13**, 527–542.

Carmona, R. and Ludkovski, M. (2004) Spot convenience yield models for the energy markets. *Contemporary Mathematics*, **351**, 65–80.

Casassus, M., Collin-Dufresne, R. and Osuntuyi, A. (2005) Stochastic convenience yield implied for commodity futures markets and interest rates. *Journal of Finance*, **60**, 2283–2331.

Chang, C. L., McAleer, M. and Tansuchat, R. (2011) Crude oil hedging strategies using dynamic multivariate GARCH. *Energy Economics*, **33**, 912–923.

Chiu, W. H. (2010) Skewness preference, risk taking and expected utility maximization. *The Geneva Risk and Insurance Review*, **35**, 108–129.

Da Fonseca, J. and Zaatour, R. (2014) Hawkes process: Fast calibration, application to trade clustering and diffusive limit. *Journal of Futures Markets*, **34**, 548–579.

Dahlquist, M., Farago, A. and Tédongap, R. (2017) Asymmetries and portfolio choice. *The Review of Financial Studies*, **30**, 667–702.

Daley, D. J. and Vere-Jones, D. (2008) *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods.* Springer.

Dassios, A. and Zhao, H. (2013) Exact simulation of Hawkes process with exponentially decaying intensity. *Electronic Communications in Probability*, **18**, 1–13.

Eraker, B. (2004) Do stock prices and volatility jump? reconciling evidence from spot and option prices. *Journal of Finance*, **59**, 1367–1403.

Eraker, B., Johannes, M. and Polson, N. (2003) The impact of jumps in volatility and returns. *Journal of Finance*, **58**, 1269–1300.

Fileccia, G. and Sgarra, C. (2015) Historical and risk-neutral estimation in a two factor stochastic volatility model for oil market. *International Journal of Computational Economics and Econometrics*, **5**, 451–479.

Fileccia, G. and Sgarra, C. (2018) A particle filtering approach to oil futures price calibration and forecasting. *Journal of Commodity Markets*, **9**, 21–34.

Filimonov, V., Bicchetti, D., Maystre, N. and Sornette, D. (2014) Quantification of the high level of endogeneity and of structural regime shifts in commodity markets. *Journal of International Money and Finance*, **42**, 174–192.

Fulop, A. and Li, J. (2013) Efficient learning via simulation: a marginalized resample-move approach. *Journal of Econometrics*, **176**, 146–161.

Fulop, A. and Li, J. (2019) Bayesian estimation of dynamic asset pricing models with informative observations. *Journal of Econometrics*, **209**, 114–138.

Fulop, A., Li, J. and Yu, J. (2015) Self-exciting jumps, learning, and asset pricing implications. *The Review of Financial Studies*, **28**, 876–912.

Gibson, R. and Schwartz, E. S. (1990) Stochastic convenience yield and the pricing of oil contingent claims. *Journal of Finance*, **45**, 959–976.

Gordon, N. J., Salmond, D. J. and Smith, A. F. M. (1993) A novel approach to nonlinear and non gaussian bayesian state estimation. *IEEE Proceedings. Part F: Radar and Sonar Navigation*, **140**, 107–113.

Ioannidis, C. and Ka, K. (2018) The impact of oil price shocks on the term structure of interest rates. *Energy Economics*, **72**, 601–620.

Kilian, L. and Figfusson, R. (2013) Do oil prices help forecast U.S. real GDP? the role of nonlinearities and asymmetries. *Journal of Business and Economic Statistics*, **31**, 78–93.

Kraus, A. and Litzenberger, R. (1976) Skewness preferences and valuation of risk assets. *Journal of Finance*, **31**, 1085–1100.

Kroner, K. F. and Sultan, J. (1993) Time-varying distributions and dynamic hedging with foreign currency futures. *Journal of Financial and Quantitative Analysis*, **28**, 531–551.

Lai, A. N. and Mellios, C. (2016) Valuation of commodity derivatives with an unobservable convenience yield. *Computers and Operations Research*, **66**, 402–414.

Larsson, K. and Nossman, N. (2011) Jumps and stochastic volatility in oil prices: Time series evidence. *Energy Economics*, **33**, 504–514.

Lautier, D. (2009) Convenience yield and commodity markets. *Bankers, Markets and Investors*, **102**, 59–66.

Lewis, P. A. and Shedler, G. S. (1969) Simulation of nonhonmogeneous poisson processes by thinning. *Naval Research Logistic Quarterly*, **26.3**, 403–413.

Litzenberger, R. and Rabinowitz, N. (1995) Backwardation in oil futures market: Theory and empirical evidence. *Journal of Finance*, **50**, 1517–1545.

Liu, Q., Chng, T. and Xu, D. (2014) Hedging industrial metals with stochastic volatility models. *Journal of Futures Markets*, **34**, 704–730.

Ma, F., Zhang, Y., Huang, D. and Lai, X. (2018) Forecasting oil futures price volatility: New evidence from realized range-based volatility. *Energy Economics*, **75**, 400–409.

Maneesoonthorn, W., Forbes, C. S. and Martin, G. M. (2016) Inference on self-exciting jumps in prices and volatility using high frequency measures. *Journal of Applied Econometrics*, **32**, 504–532.

Ogata, Y. (1981) On the Lewis' simulation method for point processes. *IEEE Transactions on Information Theory*, **27.1**, 23–31.

Pan, J. (2002) The jump-risk premia implicit in option: Evidence from an integrated time series study. *Journal of Financial Economics*, **63**, 3–50.

Post, T., Vliet, P. and Levy, H. (2008) Risk aversion and skewness preference. *Journal of Banking and Finance*, **32**, 1178–1187.

Ribeiro, D. and Hodges, S. D. (2004) A two factor model for commodity prices and futures valuation. *EFMA 2004 Basel Meetings Paper. Available at SSRN: https://ssrn.com/abstract=498802 or http://dx.doi.org/10.2139/ssrn.498802.*

Schwartz, E. S. (1997) The stochastic behavior of commodity prices: Implications for valuation and hedging. *Journal of Finance*, **52**, 923–973.

Wang, Y. (2013) Oil price effects on personal consumption expenditures. *Energy Economics*, **36**, 198–204.

Yan, X. (2002) Valuation of commodity derivatives in a new multi-factor model. *Review of Derivative Research*, **5**, 251–271.

Yu, C. L., Li, H. and Wells, M. T. (2011) MCMC estimation of Levy jump models using stock and option prices. *Mathematical Finance*, **21**, 383–422.

# A An Hybrid Particle Filter

Our estimation strategy is based on Bayesian inference. Hence, if we denote the set of model parameters as $\Theta$ and all observations and latent states as $y_{1:T} = \{\ln S_t, \ln F(t,T)^O\}_{t=1}^T$ and $x_{1:T} = \{V_t, \delta_t, \lambda_t\}_{t=1}^T$ respectively, we can define the joint posterior distribution as

$$p(\Theta, x_{1:T} \mid y_{1:T}) = p(x_{1:T} \mid \Theta, y_{1:T}) p(\Theta \mid y_{1:T}) \tag{34}$$

The state-space model we are concerned with is clearly non-linear and non-Gaussian, thus standard Kalman filtering techniques cannot be applied. Therefore, we rely on the application of particle filtering methods, which are simulation based methods able to take into account the complexity of our model. The general idea is to approximate continuous time distributions with discrete points, called particles.

Given a set of particles $\{x_{t-1}^{(i)}\}_{i=1}^M$ representing the filtering distribution $p(X_{t-1} \mid y_{1:t-1})$ at time $t-1$, we can decompose the filtering distribution at time $t$ as follows:

$$p(x_t \mid y_{1:t}) \propto \int p(y_t \mid x_t) p(x_t \mid x_{t-1}) p(x_{t-1} \mid y_{1:t-1}) dx_{t-1}. \tag{35}$$

Now, by importance sampling we can sample from a proposal density $q(x_t \mid x_{t-1})$ and then attach importance weights $w_t$ to account for the difference between the target and the proposal:

$$w_t^{(i)} = \frac{p(y_t \mid x_t^{(i)}) p(x_t^{(i)} \mid x_{t-1}^{(i)})}{q(x_t^{(i)} \mid x_{t-1}^{(i)})}, \tag{36}$$

and denote with $W_t^{(i)} = w_t^{(i)} / \sum_{j=1}^M w_t^{(j)}$ the normalized weights. This algorithm is known as sequential importance sampling and suffers from the so called *particle degeneracy*. To solve this problem we resample the particles proportional to $W_t^{(i)}$, obtaining an equally weighted sample that can be used to approximate the filtering distribution:

$$\hat{p}(x_t \mid y_{1:t}) = \frac{1}{M} \sum_{i=1}^M \delta_{x_t^{(i)}}(\tilde{x}_t^{(i)}), \tag{37}$$

where, $\tilde{x}_t^{(i)}$ are the resampled particles. This is known in the literature as sequential importance resampling (SIR). As a byproduct, the particle filter algorithm delivers an unbiased estimate of

the marginal likelihood:

$$\hat{p}(y_{1:t} \mid \Theta) = \prod_{l=2}^{T} \hat{p}(y_l \mid y_{1:l-1}, \Theta)\hat{p}(y_1 \mid \Theta), \tag{38}$$

where,

$$\hat{p}(y_l \mid y_{1:l-1}, \Theta) = \frac{1}{M}\sum_{i=1}^{M} w_l^{(i)}. \tag{39}$$

The most common particle filter is the bootstrap filter presented in Gordon *et al.* (1993), where the proposal density is simply the state transition law, i.e. $q(x_t \mid x_{t-1}) = p(x_t \mid x_{t-1})$, which does not take into account the new observation $y_t$, leading to poor performances if the observation is highly informative (e.g. in the case of a jump). In our setting we are not able to derive the optimal proposal, but we can design an hybrid smooth SIR, which is a relatively efficient particle filter with respect to outliers. In particular, to robustify the filter against outliers, we propose jump times $\Delta N_t$ from a Bernoulli with probability 0.5. Furthermore (as in Fulop and Li, 2019) instead of resampling, we fit a multivariate normal distribution on the smoothing distribution and sample from this normal using the inverse CDF method. The detailed algorithm is outlined below.

At time $t$ we have from time $t-1$ a weighted sample of $M$ particles representing the filtering distribution: $x_{t-1}^{(i)} = \{V_{t-1}^{(i)}, \lambda_{t-1}^{(i)}, \delta_{t-1}^{(i)}\}_{i=1}^{M}$.

Hence, for each particle, the PF works as follows:

- **Step 1 (Sample $\Delta N_t^{(i)}$ from proposal):** $\Delta N_t^{(i)} \sim Ber(0.5)$. Then, sample return jumps $J_x^{(i)} \sim \mathcal{N}(\mu_J, \sigma_J^2)$, variance jumps $J_v^{(i)} \sim Exp(\mu_v)$ and the latent states from their transitions:

$$\begin{aligned}
\lambda_t^{(i)} &= \lambda_{t-1}^{(i)} + \beta(\lambda_\infty - \lambda_{t-1}^{(i)})\Delta t + \alpha \Delta N_t^{(i)} \\
V_t^{(i)} &= V_{t-1}^{(i)} + k(\bar{V} - V_{t-1}^{(i)})\Delta t + \sigma_v \sqrt{\Delta t V_{t-1}^{(i)}} \left(\rho_v Z_t^{(i)} + \sqrt{1-\rho_v^2} W_{v,t}^{(i)}\right) + J_v^{(i)} \Delta N_t^{(i)} \\
\delta_t^{(i)} &= \delta_{t-1}^{(i)} + \gamma(\bar{\delta} - \delta_{t-1}^{(i)})\Delta t + \sigma_\delta \sqrt{\Delta t} W_{\delta,t}^{(i)}
\end{aligned}$$

where, $W_{v,t}^{(i)}$ and $W_{\delta,t}^{(i)}$ are independent normals and

$$Z_t^{(i)} = \frac{\ln S_t - \ln S_{t-1} - \mu_x \Delta t - J_x^{(i)} \Delta N_t^{(i)}}{V_{t-1}^{(i)} \Delta t},$$

with, $\mu_x = \mu - \frac{1}{2}V_{t-1} - \lambda_{t-1}\mu^\star - \delta_{t-1}$.

- **Step 2 (Reweight):** Compute log-weights according to

$$\log p(\ln S_t \mid \ln S_{t-1}, x_{t-1}^{(i)}, \Delta N_t) + \log p(\ln F_{t,T} \mid \ln S_t, \delta_{t-1}^{(i)}) + \log \pi_t \to \log w_t^{(i)}$$

where,

$$\begin{aligned} \text{if} \quad \Delta N_t^{(i)} = 0 \quad &\Rightarrow \quad \log \pi_t = -\lambda_t^{(i)} \Delta t - \log(0.5) \\ \text{if} \quad \Delta N_t^{(i)} = 1 \quad &\Rightarrow \quad \log \pi_t = -\log(1 - \exp(\lambda_t^{(i)} \Delta t)) - \log(0.5) \end{aligned}$$

- **Step 3 (Smooth approximation):** Generate from $x_{t-1|t} = \{V_{t-1|t}, \delta_{t-1|t}, \lambda_{t-1|t}\}$ fitting a multivariate normal. As suggested by Fulop and Li (2019) we add a moment-matching step.

# B Optimal Hedging Solution

In this appendix we provide detailed calculations of the optimal hedge ratio. Let us start from the portfolio variance:

$$\frac{1}{dt}\mathrm{Var}_t[d\Pi_t] = \frac{1}{dt}\left(\mathrm{Var}_t[dX_t] + h^2\mathrm{Var}_t[dY_t] - 2h\mathrm{Cov}_t[dX_t, dY_t]\right).$$

consider the first term:

$$\begin{aligned} \frac{1}{dt}\mathrm{Var}_t[dX_t] =& \frac{1}{dt}\mathrm{Var}_t\left[\left(r - \frac{1}{2}V_t t - \delta_t\right)dt + \sqrt{V_t}dW_{x,t} + dJ_{x,t}\right], \\ =& V_t + \frac{1}{dt}\mathrm{Var}_t\left[J_x dN_t\right], \end{aligned}$$

where,

$$\begin{aligned} \mathrm{Var}_t\left[J_x dN_t\right] =& \mathbb{E}[J_x^2 dN_t^2] - \left(\mathbb{E}[J_x dN_t]\right)^2 = \mathbb{E}[J_x^2]\mathbb{E}[dN_t] - \left(\mathbb{E}[J_x]\mathbb{E}[dN_t]\right)^2, \\ =& (\mu_J^2 + \sigma_J^2)\lambda_t dt - \mu_J \lambda_t^2 dt^2 = (\mu_J^2 + \sigma_J^2)\lambda_t dt. \end{aligned}$$

Therefore we have

$$\frac{1}{dt}\mathrm{Var}_t[dX_t] = V_t + (\mu_J^2 + \sigma_J^2)\lambda_t.$$

The next object is the instantaneous variance of log-futures:

$$\begin{aligned}\frac{1}{dt}\mathrm{Var}_t[dY_t] =& \frac{1}{dt}\mathrm{Var}_t\left[\sqrt{V_t}dW_{x,t} + C(\tau)\sigma_\delta dW_{\delta,t} + J_x dN_t\right], \\ &= V_t + C^2(\tau)\sigma_\delta^2 + (\mu_J^2 + \sigma_J^2)\lambda_t.\end{aligned}$$

The covariance between $X$ and $Y$ is given by

$$\begin{aligned}\frac{1}{dt}\mathrm{Cov}_t[dX_t, dY_t] &= \frac{1}{dt}\mathrm{Cov}_t\Big[\sqrt{V_t}dW_{x,t} + J_x dN_t, \sqrt{V_t}dW_{x,t} + C(\tau)\sigma_\delta dW_{\delta,t} + J_x dN_t\Big], \\ &= V_t + (\mu_J^2 + \sigma_J^2)\lambda_t.\end{aligned}$$

Then we obtained the conditional variance of the portfolio:

$$\frac{1}{dt}\mathrm{Var}_t[d\Pi_t] = V_t + (\mu_J^2 + \sigma_J^2)\lambda_t + h^2\Big[V_t + C^2(\tau)\sigma_\delta^2 + (\mu_J^2 + \sigma_J^2)\lambda_t\Big] - 2h\Big[V_t + (\mu_J^2 + \sigma_J^2)\lambda_t\Big].$$

Now we consider the (non-standardized) third moment:

$$\frac{1}{dt}\mathbb{E}_t\left[(d\Pi - \mathbb{E}(d\Pi))^3\right] = \frac{1}{dt}\mathbb{E}_t\left[d\Pi^3 - (\mathbb{E}(d\Pi))^3 - 3d\Pi^2\mathbb{E}(d\Pi) + 3d\Pi(\mathbb{E}(d\Pi))^2\right].$$

The first element can be computed as follows

$$\begin{aligned}\frac{1}{dt}\mathbb{E}_t\left[d\Pi^3\right] &= \frac{1}{dt}\mathbb{E}_t\left[dX^3 - h^3dY^3 - 3hdX^2dY + 3h^2dXdY^2\right], \\ &= (\mu_J^3 + 3\mu_J\sigma_J^2)\lambda_t - h^3(\mu_J^3 + 3\mu_J\sigma_J^2)\lambda_t - 3h(\mu_J^3 + 3\mu_J\sigma_J^2)\lambda_t + 3h^2(\mu_J^3 + 3\mu_J\sigma_J^2)\lambda_t.\end{aligned}$$

Now define,

$$\begin{aligned}\mu_X &= r - \frac{1}{2}V_t - \delta_t, \\ \mu_Y &= r - \frac{1}{2}V_t - \delta_t - A'(\tau) - C'(\tau)\delta_t + C(\tau)\gamma(\bar{\delta} - \delta_t).\end{aligned}$$

Then we compute,

$$\begin{aligned}\mathbb{E}_t\left[d\Pi\right] &= \mathbb{E}_t\left[dX\right] - h\mathbb{E}_t\left[dY\right], \\ &= \mu_X dt + \mu_J\lambda_t dt - h\left(\mu_Y dt + \mu_J\lambda_t dt\right).\end{aligned}$$

given this we have that $\left(\mathbb{E}_t\left[d\Pi\right]\right)^3 = \left(\mathbb{E}_t\left[d\Pi\right]\right)^2 = \mathbb{E}_t\left[d\Pi^2\mathbb{E}_t\left[d\Pi\right]\right] = 0$. Therefore,

$$\begin{aligned}&\frac{1}{dt}\mathbb{E}_t\left[\left(d\Pi - \mathbb{E}(d\Pi)\right)^3\right] = \frac{1}{dt}\mathbb{E}_t\left[d\Pi^3\right], \\ &= (\mu_J^3 + 3\mu_J\sigma_J^2)\lambda_t - h^3(\mu_J^3 + 3\mu_J\sigma_J^2)\lambda_t - 3h(\mu_J^3 + 3\mu_J\sigma_J^2)\lambda_t + 3h^2(\mu_J^3 + 3\mu_J\sigma_J^2)\lambda_t.\end{aligned}$$

Now, we take the derivative of the objective function with respect tot h,

$$\begin{aligned}\mathrm{Var}_t(\Pi_t) - \eta\mathrm{Asy}_t(\Pi_t) = V_t + (\mu_J^2 + \sigma_J^2)\lambda_t + h^2\left[V_t + C^2(\tau)\sigma_\delta^2 + (\mu_J^2 + \sigma_J^2)\lambda_t\right] - 2h\left[V_t + (\mu_J^2 + \sigma_J^2)\lambda_t\right] \\ - \eta\left[(\mu_J^3 + 3\mu_J\sigma_J^2)\lambda_t - h^3(\mu_J^3 + 3\mu_J\sigma_J^2)\lambda_t - 3h(\mu_J^3 + 3\mu_J\sigma_J^2)\lambda_t + 3h^2(\mu_J^3 + 3\mu_J\sigma_J^2)\lambda_t\right].\end{aligned}$$

If we set this quantity equal to zero and solve for h we get (31).