Journal of Econometrics and Statistics
Vol. 5, Issue 1, 2025, pp. 57-80

Parameter Estimation in Levy Driven Stochastic Volatility Models

Jaya P. N. Bishwal

Department of Mathematics and Statistics, University of North Carolina at Charlotte, 376 Fretwell Buildings, 9201 University City Boulevard, Charlotte, NC 28223-0001, USA

ABSTRACT
We generalize Ornstein-Uhlenbeck process to include non-normal innovations. This model captures the stylized facts of financial markets as it preserves jumps in the volatility process. We study the asymptotic behavior of some estimators of the drift parameter in the Gamma-Ornstein-Uhlenbeck, Inverse-Gaussian-Ornstein-Uhlenbeck, Modified-Tempered-Stable-Ornstein-Uhlenbeck volatility processes based on discrete equally spaced observations of the price process. The estimators are explicit. We study robustness and efficiency of the estimators.

KEYWORDS
Stochastic differential equation, Stochastic volatility, Levy process, Inverse Gaussian-Ornstein-Uhlenbeck process, Gamma-Ornstein-Uhlenbeck process, Modified Tempered Stable-Ornstein-Uhlenbeck process, Jumps, Infinite Divisibility, Heavy Tails, Skewness, Kurtosis, Discrete Observations, Robust Estimator, Moment Estimator, Mixing, Factor Model.

1. Introduction

Recently processes with jumps and long memory have received attention in finance, engineering and physics. Levy driven processes of Ornstein-Uhlenbeck type have been extensively studied over the last few years and widely used in finance, see Barndorff-Neilsen and Shephard [2, 3]. Levy processes are processes with stationary independent increments. Levy Ornstein-Uhlenbeck (LOU) process generalizes the Ornstein-Uhlenbeck process to include jumps.

The Levy Ornstein-Uhlenbeck (LOU) process, is an extension of Ornstein-Uhlenbeck process with Levy process driving term. In finance, it is useful as a generalization of Vasicek model, as one-factor short-term interest rate model which could take into account the jump of the interest rate. It also generalizes stochastic volatility model where the volatility has jumps.

Jump processes are of two types: Finite activity processes and infinite activity processes. Finite activity processes have finite number of jumps in a finite time interval, e.g., a Poisson process and infinite activity processes have infinite number of jumps in a finite time interval, e.g., gamma process, inverse Gaussian process and tempered stable process. We will consider only infinite activity process in this paper.

CONTACT Jaya P. N. Bishwal. Email: J.Bishwal@charlotte.edu

Article History
Received : 28 August 2024; Revised : 24 September 2024; Accepted : 12 November 2024; Published : 08 January 2025

To cite this paper

Jaya P. N. Bishwal (2025). Parameter Estimation in Levy Driven Stochastic Volatility Models. *Journal of Econometrics and Statistics*. 5(1), 57-80.

It is well known that the suitably parametrized autoregresive (AR) process with Gaussian error has the continuous limit the Vasicek model. Wolfe [56] studied continuous analogue of the stochastic difference equation of AR type with Levy type innovations whose limit is a Levy driven OU Process. Gourieroux and Jasiak [36] studied autoregressive gamma (ARG) process and showed that its continuous time limit is the Cox-Ingersoll-Ross (CIR) model. Thus the stationary ARG process is a discretized version of the CIR process. Gourieroux and Jasiak [36] studied pseudo-maximum likelihood estimation in autoregressive gamma (ARG) process. This process can also be used for application in series of squared returns and intertrade durations for high-frequency data, i.e., it is a stochastic duration model. ARG model also fits a series of volumes per trade, which is an alternative proxy for liquidity. This is different from gamma autoregressive process (GAR) process studied in Sim [54] and Gaver and Lewis [34] where just the noise of the linear autoregressive process is Gamma distributed. For intertrade durations, the most popular model is autoregressive conditional duration (ACD) model introduced by Engle and Russell [30].

Based on direct observations from the model, Bishwal [7] studied estimation by estimating function for discretely sampled diffusions. Bishwal [9] studied M-estimation for discretely sampled diffusions. Bishwal [11] studied estimation by sequential Monte Carlo method for stochastic volatility models. Bishwal [12] studied sufficiency problem in Vasicek model. Bishwal [18] studied nonparametric estimation in Heath-Jarrow-Morton forward interest model driven by Levy process using local time. Bishwal [19] studied higher order approximate maximum likelihood estimation for CKLS model.

In order to do pricing of options for these semi-observed models, the unknown parameters in the hidden model must be estimated from the asset price data. Bishwal [13] studied quasi-maximum likelihood estimation in fractional Levy driven Ornstein-Uhlenbeck stochastic volatility model.

In this paper we study method of moments estimators for Ornstein-Uhlenbeck stochastic volatility model driven by Levy processes. In section 2, we review option pricing in stochastic interest rate and stochastic volatility models, and introduce a hybrid model with 15 parameters. In section 3, we study estimation for the Inverse-Gaussian-Ornstein-Uhlenbeck stochastic volatility model. In Section 4, Gamma-Ornstein-Uhlenbeck stochastic volatility model. In section 5, we study estimation in Modified-Tempered-Stable-Ornstein-Uhlenbeck stochastic volatility model. In section 6, we make concluding remarks.

2. Option Pricing in Stochastic Interest Rate and Stochastic Volatility Models

First we consider the stock price modeling. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a stochastic basis on which is defined the following process $\{S_t, t \geq 0\}$ where $\{W_t\}$ be a standard Wiener process with the filtration $\{\mathcal{F}_t\}_{t\geq 0}$.

For the Black-Scholes model for stock price

$$dS_t = \mu S_t dt + \sigma S_t dW_t,\ t \geq 0,$$

using Itô formula to $\log S_t$, one obtains the solution

$$S_t = S_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma W_t\right)$$

which is known as *geometric Brownian motion.* The parameter μ is known as the mean rate of return and σ as the volatility.

Call option (buyer's option) at time t is the expected discounted (at the risk free interest rate r pay-off

$$C_t = E[e^{-r(T-t)}\max(S_T - K, 0)|\mathcal{F}_t]$$

where K is the strike price of the option and T is the time of maturity of the option.

Put option (seller's option) at time t is the expected discounted (at the risk free interest rate r pay-off

$$P_t = E[e^{-r(T-t)}\max(K - S_T, 0)|\mathcal{F}_t]$$

where K is the strike price of the option and T is the time of maturity of the option

Using no arbitrage principle, Black-Scholes derived the partial differential equation (PDE) which is given by

$$\frac{\partial C_t}{\partial t} + rS_t\frac{\partial C_t}{\partial S_t} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 C_t}{\partial^2 S_t} - rS_t = 0.$$

Black and Scholes [21] calculated the above expectation by solving the PDE for C_t and is known as the famous Black-Scholes option pricing formula.

The Black-Scholes option price formulae for European call and put options are given respectively by

$$C_t = E[e^{-r(T-t)}(S_T - K)^+|\mathcal{F}_t] = S_t\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2),$$

$$P_t = E[e^{-r(T-t)}(K - S_T)^+|\mathcal{F}_t] = Ke^{-r(T-t)}\Phi(-d_2) - S_t\Phi(-d_1)$$

where

$$d_1 := \frac{\log\left(\frac{S}{K}\right) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}, \quad d_2 := \frac{\log\left(\frac{S}{K}\right) + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} = d_1 - \sigma\sqrt{T-t}$$

and Φ is the cummulative distribution of standard normal distribution.

In a risk neutral world, where all expectations are calculated under the risk-neutral measure or the martingale measure, the stock price S_t at time t follows the following linear Itô stochastic differential equation, known as the Black-Scholes model

$$dS_t = rS_t dt + \sigma S_t dW_t, \ t \geq 0$$

where $\{W_t\}_{t\geq 0}$ is a standard Brownian motion, r is the risk-free interest rate and σ is the volatility. A simple application of Itô's formula to $\log S_t$ provides the exact solution of the equation given by

$$S_t = S_0\exp\{(r - \frac{1}{2}\sigma^2)t + \sigma W_t\}$$

where S_0 is the initial price of the stock. S_t is called Geometric Brownian motion.

One can generates the path of the stock price by the exact method and the Euler method and calculate the price of the European call option at time 0 based on both the methods.

In the Black-Scholes model, the interest rate r and the volatility σ are constant. In practice, both interest rates and the volatility are stochastic processes. First, we consider stochastic interest rate, which is known as short rate. The Vasicek model for short rate is given by

$$dr_t = a(b - r_t)dt + \sigma dW_t,\ t \geq 0.$$

The price of a zero coupon bond at time t maturing at time T is given by

$$P(t,T) = A(t,T)e^{-B(t,T)r(t)}$$

where

$$B(t,T) := \frac{1 - e^{-a(T-t)}}{a},\ A(t,T) := \exp\left(\frac{(B(t,T) - T + t)(a^2 b - \sigma^2/2)}{a^2} - \frac{\sigma^2 B(t,T)^2}{4a}\right).$$

The interest rate derivative, European call option is given by

$$C = LP(0,s)\Phi(h) - KP(0,T)\Phi(h - \sigma_p)$$

where L is the bond principal, s is the bond maturity, T is the option maturity, K is the strike price,

$$h := \frac{1}{\sigma_P} \ln \frac{LP(0,s)}{P(0,T)K} + \frac{\sigma_P}{2},\ \text{where } \sigma_P := \frac{\sigma}{a}(1 - e^{-a(s-T)})\sqrt{\frac{1 - e^{-2aT}}{2a}}.$$

When $a = 0$, $\sigma_P = \sigma(s - T)\sqrt{T}$. Parameter estimation in Vasicek model is extensively studied in Bishwal [8].

The drawback of the Vasicek model is that interest rate can be negative since the transition density of the process is normal. Next, we consider a positive interest rate model which also serves as stochastic volatility model. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a stochastic basis on which is defined the Cox-Ingersoll-Ross process $\{\sigma_t^2\}$ satisfying the Itô stochastic differential equation

$$d\sigma_t^2 = (1 + 2\theta\ \sigma_t^2)\ dt + \ 2\sigma_t\ dW_t,\ t \geq 0,$$

where $\{W_t\}_{t\geq 0}$ is a standard Wiener process with the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ and consider the classical direct estimation problem where $\theta < 0$ is the unknown parameter to be estimated on the basis of discrete observations of the process $\{\sigma_t^2\}_{t\geq 0}$ at times $0 = t_0 < t_1 < \cdots t_n = T$ with $t_i - t_{i-1} = \frac{T}{n} = \Delta, i = 1, 2\cdots, n$. For our asymptotic framework, we consider two types of data: 1) $\Delta \to 0, n \to \infty$ (high frequency case), 2) Δ fixed, $n \to \infty$ (low frequency case).

For the moment, assume that we have a continuous realization $\{\sigma_t^2, 0 \leq t \leq T\}$ which is denoted by ${\sigma^2}_0^T$. Let P_θ^T be the measure generated on the space (C_T, B_T) of continuous functions on $[0,T]$ with the associated Borel σ-algebra B_T generated under the supremum norm by the process V_0^T and let P_0^T be the standard Wiener measure.

It is well known that when θ is the true value of the parameter P_θ^T is absolutely continuous with respect to P_0^T and the Radon-Nikodym derivative (likelihood) of P_θ^T with respect to P_0^T based on ${\sigma^2}_0^T$ is given by

$$L_T(\theta) := \frac{dP_\theta^T}{dP_0^T}({\sigma^2}_0^T) = \exp\left\{\theta \int_0^T d\sigma_t^2 - \frac{\theta^2}{2}\int_0^T \sigma_t^2 dt\right\}.$$

Consider the score function, the derivative of the log-likelihood function, which is given by

$$\gamma_T(\theta) := \int_0^T d\sigma_t^2 - \theta \int_0^T \sigma_t^2 dt.$$

A solution of the estimating equation $\gamma_T(\theta) = 0$ provides the maximum likelihood estimate (MLE)

$$\hat{\theta}_T := \frac{\sigma_T^2 - \sigma_0^2 - T/2}{\int_0^T \sigma_t^2 dt}.$$

The minimum contrast estimate (MCE) is given by

$$\widetilde{\theta}_T := -\frac{T}{2\int_0^T \sigma_t^2 dt}.$$

Note that the volatility which is given by the CIR process is not observed.
Consider the *Heston stochastic volatility model*

$$dS_t = \mu\ S_t\ dt + \sqrt{X_t}\ S_t\ \ dW_t, \quad dX_t = (1 + 2\theta\ X_t)\ dt + \ 2\sqrt{X_t}\ \ dZ_t$$

where $\{W_t\}$ is a standard Brownian motion independent of another standard Brownian motion $\{Z_t\}$ and $\theta < 0$. The integrated volatility is given by $I_T := \int_0^T X_t dt$. Denote $s_t = \ln S_t$, $\Delta s_{t_{i-1}}^2 := (s_{t_i} - s_{t_{i-1}})^2$. The realized volatility is defined as

$$R_{n,T} := \sum_{i=1}^n \Delta s_{t_{i-1}}^2.$$

It is well known that $\operatorname*{P-lim}_{n\to\infty} R_{n,T} = I_T$.

Thus the realized volatility estimates the integrated volatility. Bishwal [16] obtained several higher order new estimators of integrated volatility using kernel method.

Note that the volatility which is given by the CIR process is not observed. In the following, we obtain nonparametric estimators of the minimum contrast estimator of the mean reversion parameter in the Heston model using approximations to $\widetilde{\theta}_T$.

Hence from the definition of MCE, the approximate minimum contrast estimate (AMCE) of θ would be

$$\widetilde{\theta}_{n,T} := -\frac{T}{2R_{n,T}}.$$

The following characteristic function of I_T is closely associated with *Levy's stochastic area formula* and is well known from Brownian motion literature and also from the work of Cox, Ingersoll and Ross [26]. Consider the special CIR model

$$dX_t = (1 + 2\theta X_t)\ dt + \ 2\sqrt{X_t}\ dW_t,\ t \geq 0.$$

Let $\phi_T(u) := E\exp(iuI_T), u \in \ \mathbb{R}$ be the characteristic function of I_T. Then

$$E\exp(iuI_T) = \exp\left(\frac{2iu}{2\theta + \gamma\coth\frac{\gamma T}{2}}\right)\left[\cosh\frac{\gamma T}{2} + \frac{2\theta}{\gamma}\sinh\frac{\gamma T}{2}\right]^{-1}$$

where $\gamma := (4\theta^2 - 2iu)^{1/2}$ and we choose the principal branch of the square root.

Consider the general CIR model

$$dX_t = (a - bX_t)dt + \sqrt{2\sigma X_t}dW_t$$

where $X_0 = x > 0,\ a > 0,\ b \in \mathbb{R},\ \sigma > 0.$

Let $J_T := \int_0^T X_t^{-1}dt$ be the integrated inverse volatility. The bond price is the moment generating function of I_T while the maximum likelihood estimators (MLEs) of the parameters (a, b) are functions of X_T, I_T and J_T. The joint Laplace transform of (X_T, I_T) is given by

$$\begin{aligned} E\exp(-uX_T - vI_T) &= \left(\frac{2\rho e^{(b-\rho)T/2}}{2\sigma u(1-e^{-\rho T}) + (\rho - b)e^{-\rho T} + (\rho + b)}\right)^{a/\sigma} \\ &\times \exp\left(\frac{u((\rho + b)e^{-\rho T} + (\rho - b) + 2v(1 - e^{-\rho T})}{2\sigma u(1-e^{-\rho T}) + (\rho - b)e^{-\rho T} + (\rho + b)}\right) \end{aligned}$$

where $\rho =: \sqrt{b^2 + 4\sigma v}$.

Finally,

$$\begin{aligned} &E\exp(-uJ_T) \\ = &\frac{\Gamma(k + \frac{\nu}{2} + \frac{1}{2})}{\Gamma(\nu+1)}\left(\frac{x}{\alpha}\right)^{-k}\beta^{\frac{\nu}{2}+\frac{1}{2}}\exp\left(\frac{b}{2\sigma}\left[at - \frac{2x}{e^{bt}-1}\right]\right)\ {}_1F_1(k + \frac{\nu}{2} + \frac{1}{2}, \nu + 1, \beta) \end{aligned}$$

where

$$k =: \frac{a}{2\sigma},\ \alpha =: \frac{be^{bt}}{\sigma(e^{bt}-1)},\ \beta := \frac{bx}{\sigma(e^{bt}-1)}, \nu =: \frac{1}{\sigma}\sqrt{(\alpha - \sigma)^2 + 4u\sigma}$$

and ${}_1F_1$ is Kummer's confluent hypergeometric function. See Ben Alaya and Kebaier [4].

For the CIR model

$$dX_t = a(b - X_t)dt + \sigma\sqrt{X_t}dW_t$$

the price at time t of a zero-coupon bond that pays \$1 at time T is given by

$$P(t, T) = E_Q\left(e^{-\int_t^T X_t dt}\right) = A(t, T)e^{-B(t,T)X_t}$$

where

$$B(t,T) := \frac{2(e^{\gamma(T-t)-1})}{(\gamma+a)(e^{\gamma(T-t)-1})+2\gamma}, \quad A(t,T) := \left(\frac{2\gamma e^{(a+\gamma)(T-t)/2}}{(\gamma+a)(e^{\gamma(T-t)-1})+2\gamma}\right)^{2ab/\sigma^2}$$

where $\gamma = \sqrt{a^2 + 2\sigma^2}$ and Q is the risk-neutral measure.

The European call option is given by

$$C_t = P(t,s)\chi^2\left(2\frac{\log(A(t,s)/K)}{B(T,s)}[\phi+\psi+B(T,s)]; \frac{4ab}{\sigma^2}, \frac{2\phi^2 X_t e^{\gamma(T-t)}}{\phi+\psi+B(T,s)}\right)$$

$$-KP(t,T)\chi^2\left(2\frac{\log(A(t,s)/K)}{B(T,s)}[\phi+\psi]; \frac{4ab}{\sigma^2}, \frac{2\phi^2 X_t e^{\gamma(T-t)}}{\phi+\psi}\right)$$

where

$$\phi := \frac{2\gamma}{\sigma^2(e^{\gamma(T-t)-1})}, \quad \psi := \frac{a+\gamma}{\sigma^2}$$

and $\chi^2(x; d, \lambda)$ is the noncentral chi-square distribution with d degrees of freedom and noncentrality parameter λ.

Next consider the Heston stochastic volatility model, see Heston [37]:

$$dS_t = \mu\ S_t\ dt + \sqrt{X_t}\ S_t\ \ dW_t, \quad dX_t = (1 + 2\theta\ X_t)\ dt + \ 2\sqrt{X_t}\ \ dZ_t,\ t \geq 0$$

where $\{W_t\}_{t\geq 0}$ is a standard Brownian motion independent of another standard Brownian motion $\{Z_t\}_{t\geq 0}$ and $\theta < 0$. The integrated volatility is given by $I_T := \int_0^T X_t dt$.

Integrated volatility has to be estimated on the basis of discrete observations of the price process $\{S_t\}$ at times $0 = t_0 < t_1 < \cdots t_n = T$ with $t_i - t_{i-1} = \frac{T}{n}, i = 1, 2\cdots, n$. Let $s_t = \log S_t$ be the log-price process. Then

$$ds_t = (\mu - \frac{1}{2}X_t)dt + \sqrt{X_t}\ \ dW_t$$

Thus the drift term depends on the volatility.

Consider the Heston model with correlated noises: Under the real world measure P, we have

$$dS_t = \mu S_t dt + \sqrt{S_t} dW_t^1, \quad dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW_t^2, \quad dW_t^1 dW_t^2 = \rho dt.$$

Under the risk neutral measure Q, we have

$$dS_t = rS_t dt + \sqrt{S_t} d\widetilde{W}_t^1, \quad dV_t = \kappa^*(\theta^* - V_t)dt + \sigma\sqrt{V_t}d\widetilde{W}_t^2, \quad d\widetilde{W}_t^1 d\widetilde{W}_t^2 = \rho dt$$

where

$$\kappa^* = \kappa + \lambda, \quad \theta^* = \frac{\kappa\theta}{\kappa+\lambda}, \quad d\widetilde{W}_t^1 = dW_t^1 + \theta_t dt, \quad d\widetilde{W}_t^2 = dW_t^2 + k\sqrt{V_t}dt$$

and the Radon-Nikodym derivative is given by

$$\frac{dQ}{dP} = \exp\left\{ -\frac{1}{2}\int_0^t (\theta_s^2 + k^2 V_s) ds - \int_0^t \theta_s dW_s^1 - \int_0^t k\sqrt{V_s} dW_s^2 \right\}$$

where $\theta_t = (\mu - r)/\sqrt{V_t}$. The characteristic function of s_T under Q is given by

$$\psi_T(u) = E_Q(e^{is_T u}) = A(u,T)e^{B(u,T)}$$

where

$$\begin{aligned} A(u,T) &:= e^{iu(s_0+rT)}, \\ B(u,T) &:= \frac{-(u^2+iu)(1-e^{\gamma T})V_0}{2\gamma - (\gamma - (\kappa - \rho\sigma ui)(1-e^{\gamma T})} \\ &\quad -\frac{\kappa\theta}{\sigma^2}\left[2\log\left(\frac{2\gamma - (\gamma - \kappa - \rho\sigma ui)(1-e^{\gamma T})}{2\gamma}\right) + (\gamma - \kappa - \rho\sigma ui)T\right] \end{aligned}$$

and $\gamma =: \sqrt{(\kappa - \rho\sigma ui)^2 + (u^2 + iu)}$.

Option price is defined as $E_t^T[e^{-r(T-t)}H(T)]$ where $H(T)$ is the payoff at time T and the expectation is under the risk neutral measure. The call option with strike price K is given by

$$\widetilde{C}(S_t, V_t, t, T) = S_t P_1 - Ke^{-r(T-t)}P_2$$

where for $j = 1, 2$,

$$P_j(x, V_t, T, K) := \frac{1}{2} + \frac{1}{\pi}\int_0^\infty Re\left(\frac{e^{-iu\ln K} f_j(x, V_t, T, u)}{iu}\right) du,$$

$$x = \ln S_t, \quad f_j(x, V_t, T, u) := \exp\{C(T-t, u) + D(T-t, u)V_t + iux\},$$

$$C(T-t, u) := rui(T-t) + \frac{a}{\sigma^2}\left[(b_j - \rho\sigma ui + d)(T-t) - 2\ln\left(\frac{1 - ge^{d(T-t)}}{1-g}\right)\right],$$

$$D(T-t, u) := \frac{b_j - \rho\sigma ui + d}{\sigma^2}\left(\frac{1 - e^{d(T-t)}}{1 - ge^{d(T-t)}}\right),$$

$$g := \frac{b_j - \rho\sigma ui + d}{b_j - \rho\sigma ui - d}, \quad d := \sqrt{(\rho\sigma ui - b_j)^2 - \sigma^2(2\nu_j ui - u^2)}$$

and

$$\nu_1 = \frac{1}{2},\ \nu_2 = -\frac{1}{2},\ a = \kappa\theta,\ b_1 = \kappa + \lambda - \rho\sigma,\ b_2 = \kappa + \lambda.$$

Carr and Madan [25] used fast Fourier transform (FFT) to evaluate the integral in this option price formula. On can also use numerical quadrature, like adaptive Simpson's rule to evaluate the integral. Then matlab program can be used to calculate the option price. Alternatively, one can calculate the option price by using the Monte Carlo method after using second order discretization of the Heston model, see Glasserman [35], page 356-357.

We study parameter estimation in Heston Model

$$dS_t = rS_t dt + \sqrt{V_t} S_t dW_t^1, \quad dV_t = \theta V_t dt + \sqrt{V_t} dW_t^2, \ V_0 = \xi$$

where W_t^1 and W_t^2 are independent Brownian motions. Here S_t denotes the stock price and V_t denotes the volatility. We estimate θ based on the observations $S_{t_1}, S_{t_2}, \cdots, S_{t_n}$. Observe that

$$S_t = S_0 \exp\left\{ rt - \frac{1}{2}\int_0^t V_s ds + \int_0^t \sqrt{V_s} dW_s \right\}.$$

Introduce the process s_t defined as $s_t := \log \frac{S_t}{S_0}$. The modified observation s_t is the solution of

$$ds_t = (r - \frac{1}{2}V_t)dt + \sqrt{V_t} dW_t^1, \ s_0 = 0.$$

Hence

$$s_t^2 - 2\int_0^t s_u ds_u = \int_0^t V_u du \ \text{ and } \ \frac{d}{dt}\left(s_t^2 - 2\int_0^t s_u ds_u\right) = V_t - V_0.$$

If $\{V_t, 0 \le t \le T\}$ were observed, then the MLE of θ would have been

$$\hat{\theta}_T = \frac{V_T - V_0}{\int_0^T V_u du}.$$

Based on $\{s_t, 0 \le t \le T\}$, the MLE is given by

$$\hat{\theta}_T = \frac{\frac{d}{dT}(s_T^2 - 2\int_0^T s_u ds_u)}{s_T^2 - 2\int_0^T s_u ds_u}.$$

We have discrete data $s_{t_1}, s_{t_2}, \cdots, s_{t_n} = s_T$. The approximate MLE is given by

$$\hat{\theta}_{n,T} = \frac{\sum_{i=1}^n (s_{t_i} - s_{t_{i-1}})^2}{[s_T^2 - 2\sum_{i=1}^n s_{t_{i-1}}(s_{t_i} - s_{t_{i-1}})]\Delta}.$$

Extending the Sharpe [53]'s CAPM model, Fama and French [31, 32] studied three factor model for asset price to describe stock returns, which won all three of them Nobel prizes for their work. The three factors are 1) market risk, 2) the outperformance of small versus big companies and 3) the outperformance of high book/market versus low book/market. This model is popularly known as Fama-French model. In the last decade, Fama and French [33] introduced a five factor model for stock returns by extending the three-factor model and adding two more factors which captures the size, value, profitability and investment.

A real factor parsimonious asset price model should be of the following hybrid type with 5 factors and 15 parameters. We consider the hybrid stochastic volatility, stochastic interest rate, stochastic leverage and stochastic elasticity model under the risk neutral measure which is given by

$$dS_t = X_t dt + \sqrt{V_{t-}} S_t dW_t + \rho_{\lambda t} dL_{\tau_{\lambda t}}, \quad dV_t = -\lambda V_t dt + \upsilon_{\lambda t-} dL^H_{\tau_{\lambda t}},$$

$$dX_t = \alpha(\beta - X_t)dt + \sigma X_t^{\gamma_t} dW_t^H, \quad d\rho_t = ((2\zeta - \eta) - \eta\rho_t)dt + \theta\sqrt{(1+\rho_t)(1-\rho_t)}dZ_t,$$

$$d\xi_t = \kappa(\mu - \xi_t)dt + \varsigma\sqrt{\xi_t}dB_t, \quad d\gamma_t = \varpi(\psi - \delta))dt + \sqrt{\chi}dM_t, \quad d\tau_t = \xi_{t-}dt$$

where L_t is a Levy process, L_t^H is a fractional Levy process (see Bishwal [13, 20]), W^H is a subfractional Brownian motion, B_t, Z_t and M_t are standard Brownian motions. Here S_t is the asset price which a geometric jump-diffusion, V_t is the stochastic volatility which is a Levy Ornstein-Uhlenbeck process, X_t is the stochastic interest rate which is a sub-fractional Chan-Karloyi-Longstaff-Sanders (CKLS) process, ρ_t is the stochastic leverage Jacobi (Beta) process, ξ_t is a volatility modulation (stochastic time change) of the driving Levy subordinator which is a Cox-Ingersoll-Ross (CIR) process, γ_t is the stochastic elasticity models which is another CIR process, and all the 15 parameters $\lambda, \alpha, \beta, \sigma, \xi, \eta, \theta, \kappa, \mu, \varsigma, \varpi, \psi, \delta, \chi, H$ are positive. Estimating all the 15 parameters based on asset price and interest rate data is a challenging problem. We partially solve this problem by estimating the parameters of the unobserved volatility process based on asset price data.

3. Robust Estimation in Inverse-Gaussian-Ornstein-Uhlenbeck Stochastic Volatility Model

We consider the SDEs

$$dY_t = (\mu + \beta X_t)dt + \sqrt{X_t}dW_t + \rho dZ_{\theta t}, \quad dX_t = -\theta X_t dt + dZ_{\theta t}$$

where Z_t is a Levy process independent of X_0 with $\mathcal{L}(X_0) = IG(\delta, \gamma)$. We suppose that δ and γ are known. Here $\theta > 0$ and $\rho < 1$.

When the process Z is inverse-Gaussian, the model is IGOU process. In IGOU model, calculation of conditional cummulants of the integrated volatility conditioned on the initial value is enough to be able to compute European style options.

Note that the cumulative process or the integrated process $I_t = \int_0^t X_u du$ has long range dependence or long memory, see Barndorff-Neilsen and Shephard [1].

The cumulant functions of IGOU process are given by

$$k(u) = \log E[e^{-uZ(1)}] = -u\delta\gamma^{-1}(1 + 2u\gamma^{-2})^{-1/2},$$

$$k'(u) = \log E(e^{-uX_t}) = \delta\gamma - \delta\gamma(1 + 2u\gamma^{-2})^{1/2}, \quad u \in \mathbb{R}.$$

The process Z is the sum of two independent Levy processes $Z = Z^{(1)} + Z^{(2)}$ where $\mathcal{L}(X_0) = IG(\delta/2, \gamma)$ and $Z^{(2)}$ is a compound Poisson process given by $Z^{(2)} = \gamma^{-2}\sum_{j=1}^{N_t} u_j$ with N being a Poisson process with intensity $\delta\gamma/2$ and u_j is a sequence of independent and identically χ_1^2-distributed random variables independent of N, see Barndorff-Neilsen and Shephard [1].

The processes Z and X have infinitely many jumps in any finite time interval, hence they are infinite activity processes. Invariant distribution is Generalized Inverse Gaussian (GIG):

$$\mathcal{L}(X_0) = \mathcal{L}(X_t) = \mathcal{GIG}(\lambda, \delta, \sqrt{\alpha^2 - \beta^2}).$$

Mixture distribution is Generalized Hyperbolic (GH):

$$\mathcal{L}(Y_t) = \mathcal{GH}(\theta, \frac{\alpha}{\sqrt{t}}, \frac{\beta}{\sqrt{t}}, \sqrt{t}\delta, \mu t), \quad \mathcal{L}(Y_{t+1} - Y_t) = \mathcal{GH}(\lambda, \alpha, \beta, \delta, \mu).$$

Recall that the density function of the Generalized Hyperbolic (GH) distribution is given by

$$d_{\mathcal{GH}(\theta,\alpha,\beta,\delta,\mu)}(x) = \frac{(\alpha^2 - \beta^2)^{\frac{\theta}{2}}}{\sqrt{2\pi}\alpha^{\theta-\frac{1}{2}}\delta^\theta K_\theta(\delta\sqrt{\alpha})}(\delta^2 + (x-\mu)^2)^{(\theta-\frac{1}{2})/2} e^{\beta(x-\mu)} K_{\theta-\frac{1}{2}}(\alpha\sqrt{\delta^2 + (x-\mu)^2})$$

where K_ν denotes the modified Bessel function of third kind with index ν. When $\theta = -\frac{1}{2}$, Generalized Inverse Gaussian (GIG) becomes Inverse Gaussian $IG(\delta, \gamma)$ distribution.

First we study the LAD estimators in the stable OU case for small Δ. Consider the model

$$dX_t = -\theta X_t dt + dZ_t, \;\; t \geq 0$$

where $\{Z_t\}_{t\geq 0}$ is a Levy process independent of X_0. The Least Absolute Deviation (LAD) estimator is robust to "outlying data". The LAD estimation has a long history and is one of popular estimation procedure robust to outlers. The LAD estimation is based on the Laplacian L^1- loss while the LSE is on the Laussian L^2-loss. We refer to Portnoy and Koenker [52], Knight [41] and Koenker [42] as well as the references therein for a detailed account and historical background on LAD estimation. For time series literature, see Davis and Dunsmuir [27] and Davis, Knight and Liu [28].

The least absolute deviation (LAD) estimator is defined as the minimizer $\hat{\theta}_n$ of the contrast function $\theta \to \sum_{i=1}^n |X_{t_i} - e^{-\theta\Delta}X_{t_{i-1}}|$. For fixed T, asymptotic normality of $\hat{\theta}_n$ is achieved at the rate $n^{1/\beta-1/2}$ where β stands for the activity index of the driving Levy process, also known as the Blumenthal-Getoor [22] activity index defined as

$$\beta = \inf\left\{r > 0 : \int_{|z|\leq 1} |z|^r \nu(dz) < \infty\right\}$$

which is the degree of small-jump fluctuations.

Now consider another least absolute deviation (LAD) estimator which is defined as the minimizer $\widetilde{\theta}_n$ of the contrast function $\theta \to \sum_{i=1}^n |X_{t_i} + \theta X_{t_{i-1}}\Delta|$.

Under infill and large time sampling design, that is when $\Delta \to 0$ and $n\Delta \to \infty$, asymptotic normality of $\widetilde{\theta}_n$ is achieved at the rate $\sqrt{n}\Delta^{1-1/\beta}$ where β stands for the activity index of the driving Levy process. Note that $\sqrt{n}\Delta^{1-1/\beta} = T^{1-1/\beta}n^{(2-\beta)/(2\beta)}$. This implies that the rate of convergence is determined by the most active part of the driving Levy process, the presence of a driving Wiener part leads to $\sqrt{n\Delta}$, which is familiar in the context of asymptotically efficient estimation of diffusions with compound Poisson jumps, while a pure-jump driving Levy process leads to a faster one. As a result, when Z is a pure jump Levy process, we have a faster rate of convergence that the familiar rate $\sqrt{n\Delta}$. It is interesting to note that rate of convergence is faster by only changing the type of loss from L^2 to L^1.

Using self-weighted LAD (SLAD) contrast function $\theta \to \sum_{i=1}^{n} w(X_{t_{i-1}})|X_{t_i} + \theta X_{t_{i-1}}|$, for a bounded continuous weight function w, the rate of convergence can be improved to the conventional rate $\sqrt{n}$, see Masuda [48] which extended autoregressive process with infinite variance studied in Ling [43]. An example of weight function is $w(x) = \exp(-|x|)$. The unweighted weight function corresponds to the case $w \equiv 1$.

In the Wiener case, the unweighted SLAD estimator leans to the asymptotic variance $\pi\theta_0$ where as the asymptotic variance of the exact MLE is $2\theta_0$. Hence the asymptotic efficiency of the SALD estimator relative to the MLE is $2/\pi$.This is same as asymptotic relative efficiency for the same the sample median over sample mean in estimating the mean of i.i.d. normal samples. For the asymptotic normality of SLAD estimator one does not need the rapidly increasing experimental design $n\Delta^2 \to 0$ which is quite inevitable while adopting contrast function based on Euler-type approximation. For SLAD estimator, the weaker condition $n\Delta^3 \to 0$ is sufficient.

The SLAD estimator converges more rapidly than the LSE as soon as $\Delta = o(n^{-1/2}(\log n)^{1/(2-\beta)})$. The sampling design condition for h is $\Delta = n^{-\tau}$ where $\tau \in (0,1)$.

An interval estimator of τ is

$$\frac{\beta}{2(2\beta-1)} < \tau < \frac{\beta}{2\beta-1}, \text{ if } \beta > 1, \quad \text{and} \quad \frac{\beta}{2(2\beta-1)} < \tau < 1, \text{ if } \beta \in (2/3, 1).$$

An example of infill interval is $h = n^{-3/5}$.

Next we consider the SDEs

$$dY_t = (\mu + \beta X_t)dt + \sqrt{X_t}dW_t + \rho dZ_{\lambda t}, \quad dX_t = -\theta X_t dt + dZ_{\lambda t}$$

where Z_t is a Levy process independent of X_0 and $\mathcal{L}(X_0) = IG(\delta, \gamma)$. We suppose that δ and γ are known and we are interested in estimating $\vartheta_0 = (\theta, \rho)$, where $\theta > 0$ and $\rho < 1$.

Note that if $X_{j\Delta}, 1 \le j \le n$ were observed, then

$$\theta_n := -\frac{1}{\Delta} \ln \left(\min_{1 \le j \le n} \frac{X_{j\Delta}}{X_{(j-1)\Delta}} \right)$$

is a weakly consistent estimator of λ as $n \to \infty$ as showed by Jongbloed *et al.* [39]. In our case, $X_{j\Delta}, 1 \le j \le n$ are unobserved.

We can estimate the state process X by $\hat{X}$ given by the recursion

$$\hat{X}_{t_i} = e^{-\theta\Delta}\hat{X}_{t_{i-1}} + Z_{t_i} - Z_{t_{i-1}}.$$

Then we estimate θ by

$$\check{\theta}_n := -\frac{1}{\Delta} \ln \left(\min_{1 \le j \le n} \frac{\hat{X}_{j\Delta}}{\hat{X}_{(j-1)\Delta}} \right)$$

In order to construct the estimating functions, we use the first and second cummulants which are given respectively by

$$\kappa_{y_1}^{(1)} = \lambda \rho \Delta \kappa_{IG}^{(1)}, \quad \kappa_{y_1}^{(2)} = \Delta \kappa_{IG}^{(1)} + 2\lambda \rho^2 \Delta \kappa_{IG}^{(2)}.$$

Inverting these cummulants and replacing the cummulants by their sample quantities, we obtain the explicit the moment estimators of ρ and λ.

The moment estimators of ρ and λ are given by

$$\hat{\rho}_n := \frac{\gamma(\gamma s_y^2 - \Delta\delta)}{2\bar{y}}, \quad \hat{\theta}_n := \frac{\gamma \bar{y}}{\Delta \delta \hat{\rho}_n}$$

where

$$s_y^2 := \frac{1}{n}\sum_{j=1}^{n}(y_j - \bar{y})^2 = \frac{1}{n}\sum_{j=1}^{n} y_j^2 - (\bar{y})^2, \quad \bar{y} := \frac{1}{n}\sum_{j=1}^{n} y_j, \quad y_j := Y_{j\Delta} - Y_{(j-1)\Delta}.$$

Let $\vartheta = (\rho, \theta)$. and $\hat{\vartheta}_n = (\hat{\rho}_n, \widetilde{\lambda}_n)$. We have the following properties of the estimators from Masuda [47]:

Proposition 3.1 For fixed $\Delta > 0$ as $n \to \infty$,

$$(a) \quad \hat{\vartheta}_n \to \vartheta_0 \text{ a.s.} \quad \text{as } n \to \infty.$$

$$(b) \quad \sqrt{n}(\hat{\vartheta}_n - \vartheta_0) \to^{\mathcal{D}} \mathcal{N}_2(0, (2\lambda\rho^2\Delta^2\delta^2\gamma^{-4})^{-2} D(\vartheta_0)) \text{ as } n \to \infty.$$

where $D(\vartheta_0)$ is the limiting covariance matrix.

Remark: In the IG-OU stochastic volatility model, for the case of $\rho = -1, \ \gamma = 1, \ \delta = 1$ the moment estimator is given by $\hat{\theta}_{n\Delta} := -\frac{Y_{n\Delta} - Y_0}{n\Delta}$. Thus the parameter λ can be estimated by just the two terminal observations.

The moment estimators are sensitive to outliers since they are based on mean and standard deviation of the sample data. In order to incorporate outliers and model misspecifications, we consider robust estimators. The robust estimators of ρ and λ are given by

$$\widetilde{\rho}_n := \frac{\gamma(\gamma a_y - \Delta\delta)}{2\widetilde{y}}, \quad \widetilde{\lambda}_n := \frac{\gamma \widetilde{y}}{\Delta \delta \widetilde{\rho}_n} \quad \text{where} \quad a_y := \frac{1}{n}\sum_{j=1}^{n} |y_j - \widetilde{y}|$$

is the sample mean absolute deviation from median,

$$\widetilde{y} := \text{median of } \{y_j, 1 \le j \le n\}$$

which is defined as

$$\widetilde{y} = \begin{cases} \frac{y_k + y_{k+1}}{2} & : \quad n = 2k \\ y_{k+1} & : \quad n = 2k+1 \end{cases}$$

Let $\vartheta = (\rho, \theta)$. and $\widetilde{\vartheta}_n = (\widetilde{\rho}_n, \widetilde{\theta}_n)$. We have the following properties of the estimators. By using the standard theory of order statistics(see Theorem 5.9 and 5.21 in Van der Vaart [55]) and mixing property of the process, along with Glivenko-Cantelli argument and Delta method, we obtain:

Proposition 3.2 For fixed $\Delta > 0$ as $n \to \infty$,

$$(a) \quad \widetilde{\vartheta}_n \to \vartheta_0 \ \text{ a.s.} \quad \text{as } n \to \infty.$$

$$(b) \quad \sqrt{n}(\widetilde{\vartheta}_n - \vartheta_0) \to^{\mathcal{D}} \mathcal{N}_2(0, \frac{\pi}{2}(2\theta\rho^2\Delta^2\delta^2\gamma^{-4})^{-2} D(\vartheta_0)) \text{ as } n \to \infty.$$

where $D(\vartheta_0)$ is the limiting covariance matrix.

4. Method of Moments Estimation in Gamma-Ornstein-Uhlenbeck Stochastic Volatility Model

We generalize Ornstein-Uhlenbeck process to include non-normal innovations. First we study the asymptotic behavior of the ratio estimator of the drift parameter in Gamma-Ornstein-Uhlenbeck (GOU) volatility process based on observations of the asset price process. This model captures the stylized facts as it preserves jumps in the volatility process. We study the behavior of the moment estimators.

Bishwal [14] studied estimation for the discretely observed Ornstein-Uhlenbeck-Gamma (OUG) process. We note that this is different from Gamma-Ornstein-Uhlenbeck (GOU) process we study here. OUG process is an Ornstein-Uhlenbeck process with an additive Brownian noise and gamma noise. GOU process is a pure jump process. OUG process may be compared with the GAR process.

Recall that the autoregressive gamma (ARG) process can be written as

$$X_t = \sum_{j=1}^{N_t} U_j + \epsilon_t$$

where N_t is a Poisson process, U_j are independent and identically distributed Gamma distributed random variables with shape parameter 1 and scale parameter c_t. independent of N and ϵ_t are gamma distributed random variables with shape parameter ν and scale parameter c_t.

We propose the discretized version of the Heston model as:

$$y_i = \mu + \sqrt{x_i}\ \xi_i, \quad x_i = \sum_{j=1}^{N_t} U_{i,j} + \epsilon_i, \quad i \geq 1$$

where y_i is the return, x_i is the volatility, ξ_i are normally distributed, ϵ_i are gamma distributed, (ξ_i, ϵ_i) has the correlation ρ and U_j are independent χ_1^2 distributed.

We consider the special cases when L is an GOU process. L has the GOU density with parameters δ and γ. The stochastic volatility model has the form

$$dY_t = (\mu + \beta X_t)dt + \sqrt{X_t}dW_t + \rho dZ_t, \quad dX_t = -\theta X_t dt + dZ_t, \ t \geq 0$$

where $\{Z_t\}_{t\geq 0}$ is a Levy process independent of X_0. The invariant distribution is Gamma:

$$\mathcal{L}(X_0) = \mathcal{L}(X_t) = \mathcal{G}(\theta, \delta, \sqrt{\alpha^2 - \beta^2}).$$

Mixture distribution is Variance-Gamma denoted by Q_θ:

$$\mathcal{L}(Y_t) = \mathcal{VG}(\theta, \frac{\alpha}{\sqrt{t}}, \frac{\beta}{\sqrt{t}}, \sqrt{t}\delta, \mu t), \quad \mathcal{L}(Y_{t+1} - Y_t) = \mathcal{VG}(\theta, \alpha, \beta, \delta, \mu).$$

Let $q_t(x, \theta)$ be the density function of Y_t and $L_T(\theta)$ be the corresponding likelihood function. The contrast function is defined as $l_T(\theta) = -\log L_T(\theta)$. Minimum contrast estimator (MCE) is defined as

$$\hat{\theta}_T := \arg\inf_\theta l_T(\theta).$$

Asymptotic properties of MCE was studied in Bishwal [6, 10, 15].

Then we consider the Gamma-OU process. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a stochastic basis on which is defined the Ornstein-Uhlenbeck process X_t satisfying the Itô stochastic differential equation

$$dX_t = -\theta X_t dt + dZ_t, \quad t \geq 0,$$

where $\{Z_t\}$ is a Gamma process with the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ and $\theta > 0$ is the unknown parameter to be estimated on the basis of continuous observation of the process $\{Y_t\}$ on the time interval $[0, T]$. The solution of the above SDE is given by

$$X_t = \int_{-\infty}^{t} e^{-\theta(t-s)} dZ_s.$$

This process is stationary. In fact, it can be shown that X_{t_i} is a stationary discrete time AR (1) process with autoregression coefficient $\phi \in (0, 1)$ with the following representation

$$X_{t_i} = \phi X_{t_{i-1}} + \epsilon_{t_{i-1}}, \ i \geq 1, \text{ where } \phi = e^{-\theta\Delta}, \quad \epsilon_{t_{i-1}} = \int_{t_{i-1}}^{t_i} e^{-\theta(t_i - u)} dZ_u.$$

The moment estimators are defined as

$$\hat{\theta}_n := \frac{\frac{1}{n^2}\left[\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})\right]^2}{\frac{1}{n^2}\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})^2 - \frac{\Delta}{n}\left[\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})\right]} \frac{2a^3(a+1)}{b^4\Delta}.$$

$$\hat{\rho}_n := \frac{\frac{1}{n^2}\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})^2 - \frac{\Delta}{n}\left[\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})\right]}{\frac{1}{n^2}\left[\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})\right]} \frac{b^3\Delta}{2a^2(a+1)}.$$

For the exponential AR(1) model, the ratio estimator of θ is defined as

$$\hat{\theta}_n := -\frac{1}{\Delta}\ln\left[\min_{1\le i\le n}\frac{X_{i\Delta}}{X_{(i-1)\Delta}}\right].$$

This estimator is motivated by the extreme value theory for the correlation parameter of an AR(1) process whose innovation distribution is positive. See Davis and McCormick [29]. In the case of exponential AR(1) process, it coincides with the maximum likelihood estimator. See Neilsen and Shephard [51].

The weak consistency of the ratio estimator in the LOU process was studied in Jongbloed *et al.* [39]. The strong consistency and asymptotic Weibullness was studied in Brockwell, Davis and Yang [23] in the case of Gamma innovations.

Using the techniques of AR (1) type model with exponential innovations (Davis and McCormick [29]), we also obtain the following two estimators of the drift θ which are defined as

$$\breve{\theta}_n := -\frac{1}{\Delta}\ln\left(\min_{1\le i\le n}\frac{\hat{X}_{i\Delta}}{\hat{X}_{(i-1)\Delta}}\right), \quad \widetilde{\theta}_n := -\frac{1}{\Delta}\ln\left[\frac{\sum_{i=1}^n \hat{X}_{i\Delta}}{\sum_{i=1}^n \hat{X}_{(i-1)\Delta}}\right]$$

where $\hat{X}$ is the estimator of X based on observations of Y which could be obtained, for example, by Kitagawa algorithm. The limit distribution of the first estimator would be Weibull which can be useful for extreme value theory in finance. Bishwal [17] studied extreme value theory in finance.

Let $\vartheta = (\rho, \theta)$ and $\hat{\vartheta}_n = (\hat{\rho}_n, \hat{\theta}_n)$. By using Theorem 2.2 in Masuda [47] (see also Theorem 4.1 Van der Vaart [55]), we obtain the strong consistency and asymptotic normality of the method of moments (MM) estimators:

Proposition 4.1 For fixed $\Delta > 0$ as $n \to \infty$,

$$(a) \quad \hat{\vartheta}_n \to \vartheta_0 \text{ a.s.} \quad \text{as } n \to \infty,$$

$$(b) \quad \sqrt{n}(\hat{\vartheta}_n - \vartheta_0) \to^{\mathcal{D}} \mathcal{N}_2(0, (I^{-1}(\vartheta_0)) \text{ as } n \to \infty.$$

where $I(\vartheta_0)$ is the Fisher information matrix.

5. Method of Moments Estimation in Modified Tempered Stable-Ornstein-Uhlenbeck Stochastic Volatility Model

Masuda and Uehara [50] studied two-step estimation in ergodic Levy driven SDE

$$dX_t = a(\theta, X_t)dt + b(\beta, X_{t-})dZ_t,\ t \geq 0, \quad X_0 = x_0.$$

Masuda [49] studied multi-step estimation in stable OU Model:

$$dX_t = -\theta X_t dt + \sigma dZ_t,\ t \geq 0,\ X_0 = x_0.$$

For the least squares estimator (LSE) of θ, Hu and Long [38] obtained

$$\left(\frac{T}{\log n}\right)^{1/\beta} (\widetilde{\theta}_n - \theta_0) \to^{\mathcal{D}} 2\theta_0(\beta\theta_0)^{-1/\beta} \frac{S'_\beta}{S''^{+}_{\beta/2}}$$

as $n \to \infty$, $T \to \infty$, $T/n \to 0$, $T^{1+\beta}/n^\beta \log n \to 0$, $T^{2\beta-1}n^{2-2\beta}\log n \to \infty$, $T^{2-\beta/2+\rho}n^{-1+\beta/2-\rho} \to \infty$ for some $\rho > 0$ small enough such that all the convergence conditions are compatible, where S'_β has symmetric stable distribution of order β, $S''^{+}_{\beta/2}$ has positive stable distribution of order $\beta/2$, and S'_β and $S''^{+}_{\beta/2}$ are independent random variables.

Thus the rate at which $\widetilde{\theta}_n$ converges to θ_0 is $((\log n)/T)^{1/\beta}$, which is faster than $T^{-1/2}$ in the classical Brownian case. Also, one needs $T/n^{2/3} \to 0$ in the Brownian case, here one needs complicated design conditions for the high frequency observation sampling.

While in Gaussian OU case, for different parts $\theta > 0,\ \theta < 0$ and $\theta = 0$, LAN, LAMN and LABF hold respectively, in stable case entirely different phenomena occur.

The solution of the SDE is given by

$$X_t = e^{-\theta(t-s)}X_s + \sigma\int_s^t e^{-\theta(t-s)}dZ_u,\ t > s$$

Due to the stable integral property,

$$\mathcal{L}\left(\int_s^t e^{-\theta(t-s)}dZ_u\right) = S_\beta(\kappa_\Delta(\theta)) \text{ where } \kappa_\Delta(\theta) = \left\{\frac{1-e^{-\theta\Delta}}{\theta\beta}\right\}^{1/\beta} \sim \Delta^{1/\beta}.$$

For each $j \leq n$, the transition probability is given by

$$\mathcal{L}(X_{t_j}|X_{t_{j-1}} = x) = \delta_{x\exp(-\theta\Delta)} \star S_\beta(\kappa_\Delta(\theta)).$$

LAMN holds for $\theta \in \mathbb{R}$ when T is fixed and

$$n^{1/\beta-1/2}(\hat{\theta}_n - \theta) \to^D MN(0, \Gamma_0(T)^{-1}).$$

We study estimation in MTS-OU SV model. The IG-OU model is a special case.

An infinitely divisible distribution is said to be α-modified tampered stable distribution (α-MTS) distribution if its Levy triplet is given by

$$\sigma^2 = 0,$$
$$\nu(dx) = C\left(\frac{\lambda_+^{\alpha+\frac{1}{2}} K_{\alpha+\frac{1}{2}}\lambda_+ x}{x^{\alpha+\frac{1}{2}}} I_{x>0} + \frac{\lambda_+^{\alpha+\frac{1}{2}} K_{\alpha+\frac{1}{2}}\lambda_- x}{x^{\alpha+\frac{1}{2}}} I_{x<0}\right) dx,$$
$$\gamma = \mu + C\left(\frac{\Gamma(\frac{1}{2}-\alpha)}{2^{\alpha+\frac{1}{2}}}(\lambda_+^{2\alpha-1} - \lambda_-^{2\alpha-1}) - \lambda_+^{\alpha-\frac{1}{2}} K_{\alpha-\frac{1}{2}}(\lambda_+) + \lambda_-^{\alpha-\frac{1}{2}} K_{\alpha-\frac{1}{2}}(\lambda_-)\right)$$

where $C > 0, \lambda_+, \lambda_- > 0,\ \mu \in \mathbb{R},\ \alpha \in (-\infty, 1)\backslash\{\frac{1}{2}\}$ and $K_p(x)$ is the modified Bessel function of second kind. We denote the MTS random variable by $X \sim MTS(\alpha, C, \lambda_+, \lambda_-, \mu)$. The Levy measure $\nu(dx)$ is called the MTS Levy measure with parameter $(\alpha, C, \lambda_+, \lambda_-)$.

The MTS distribution is obtained by taking a symmetric α-stable distribution with $\alpha \in (0,1)$ and multiplying by a Levy measure with $\sqrt{|x|}\lambda^{\alpha+\frac{1}{2}} K_{\alpha+\frac{1}{2}}(\lambda|x|)$ on each half of the real axis. The measure can be extended to the case $\alpha \leq 0$. If $\alpha = \frac{1}{2}$, then γ may not be defined, so it is removed. The MTS distribution was introduced by Kim, Rachev and Chung [40].

The tails of the α-MTS distribution are thinner than those of the 2α-stable and fatter (heavier) than those of the 2α-TS distribution. At the zero neighborhood, all three have the same asymptotic behavior.

If $\lambda_+ > \lambda_-$, then the distribution is *skewed to the left.* If $\lambda_+ < \lambda_-$, then the distribution is *skewed to the right.* If $\lambda_+ = \lambda_-$, then the distribution is *symmetric.*

C controls the kurtosis of the distribution. If C increases, the peakedness of the distribution increases.

As α decreases, the distribution has fatter tails and increased peakedness. The Levy process corresponding to the MTS distribution has *finite activity* if $\alpha < 0$ and *infinite activity* if $\alpha > 0$. It has *finite variation* if $\alpha < \frac{1}{2}$ and *infinite variation* if $\alpha > \frac{1}{2}$.

With proper choice of C and μ, MTS distribution has zero mean and unit variance, and the distribution is called standard MTS distribution and denoted $X \sim stdMTS(\alpha, \lambda_+, \lambda_-)$.

CGMY process proposed in Carr *et al.* [24] is a tempered stable process. In order to obtain a closed form solution of the European option price, CGMY used the generalised Fourier transform of the distribution of the stock price under the assumption of Markov property.

The stochastic volatility model is given by

$$\begin{aligned} dY_t &= (\mu + \beta X_t)\, dt + \sqrt{X_t}\, dW_t + \rho\, dZ_t,\ t \geq 0 \\ dX_t &= -\theta\, X_t\, dt + dZ_t,\ t \geq 0 \end{aligned}$$

where μ is the drift parameter, β is the risk premium, $\theta > 0$ is the drift of the volatility and Z_t is a MTS process.

We estimate θ from the observations of $\{Y_t\}$ at the time points $t_k = k\Delta,\ k = 0, 1, 2, \ldots, n,\ \Delta > 0$. Define

$$c_m(Z) := \frac{d^m}{du^m} \log \phi_{TS}(u)|_{u=0}$$

For the tempered stable distribution $\mathrm{TS}(b, \delta, \gamma)$ where $0 < b < 1, \delta > 0, \gamma \geq 0$, the m-th cummulant is given by

$$c_m(Z) = -\delta(-2)^m \gamma^{(b-m)/b} b(b-1)\dots(b-(m-1))$$

for $\gamma > 0$. When $\gamma = 0$, it is positive b-stable distribution for which the moments of only order $k < b$ exist. For $b = 1/2$, TS distribution reduces to IG distribution.

The infinite divisibility of this distribution allows one to construct the corresponding Levy process. A Levy process $Z = (Z_t)_{t \geq 0}$ is said to be a tempered stable process if Z_1 follows a tempered stable distribution. The tempered stable process is of *finite activity* if $\alpha < 0$ and *infinite activity* if $0 < \alpha < 2$. The tempered stable process is of *finite variation* if $0 < \alpha < 1$ and *infinite variation* if $1 < \alpha < 2$.

The MTS-GARCH model is given by

$$\log \frac{S_t}{S_{t-1}} = r_t - d_t + \lambda_t \sigma_t - g(\sigma_t; \alpha, \lambda_+, \lambda_-) + \sigma_t \epsilon_t,$$

$$\sigma_t^2 = (\alpha_0 + \alpha_1 \sigma_{t-1}^2 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2) \wedge \rho, \ \epsilon_0 = 0$$

where $\alpha_0, \alpha_1, \beta_1 \geq 0, \ \alpha_1 + \beta_1 < 1, \ 0 < \rho < \lambda_+^2, \ \epsilon_t \sim stdMTS(\alpha, \lambda_+, \lambda_-)$, r_t is the risk-free rate, d_t is the dividend rate, λ_t is the market price of risk, g is the characteristic exponent of the Laplace transform for the distribution $stdMTS(\alpha, \lambda_+, \lambda_-)$, i.e., $g(x; \alpha, \lambda_+, \lambda_-) = \log(E(\exp(x\epsilon_t)))$.

The characteristic function of Z is given by

$$\phi_Z(u) = \exp(iu\mu + G_R(u; \alpha, C, \lambda_+, \lambda_-) + G_I(u; \alpha, C, \lambda_+, \lambda_-))$$

where for $u \in \mathbb{R}$,

$$\begin{aligned}
& G_R(u; \alpha, C, \lambda_+, \lambda_-) \\
= \ & 2^{-\frac{\alpha+3}{2}} \sqrt{\pi} C \Gamma\left(1 - \frac{\alpha}{2}\right) \left[(\lambda_+^2 + u^2)^{\frac{\alpha}{2}} - \lambda_+^\alpha + (\lambda_-^2 + u^2)^{\frac{\alpha}{2}} - \lambda_-^\alpha\right], \\
& G_I(u; \alpha, C, \lambda_+, \lambda_-) \\
= \ & iuC 2^{-\frac{\alpha+1}{2}} \Gamma\left(\frac{1-\alpha}{2}\right) \\
& \times \left[\lambda_+^{\alpha-1} F\left(1, \frac{1-\alpha}{2}; \frac{3}{2}, ; -\frac{u^2}{\lambda_+^2}\right) - \lambda_-^{\alpha-1} F\left(1, \frac{1-\alpha}{2}; \frac{3}{2}, ; -\frac{u^2}{\lambda_-^2}\right)\right]
\end{aligned}$$

where F is the hyper-geometric function.

The value of G_I for symmetric MTS distribution is always zero.

The m-th cumulant is given by

$$c_m(Z) = \mu \ \text{ if } \ m = 1,$$

$$c_m(Z) = 2^{m - \frac{\alpha+3}{2}} \left(\frac{m-1}{2}\right)! C \Gamma\left(\frac{m-\alpha}{2}\right) (\lambda_+^{\alpha-m} - \lambda_-^{\alpha-m}) \ \text{ if } \ m = 3, 5, 7, \dots$$

$$c_m(Z) = 2^{-\frac{\alpha+3}{2}} \sqrt{\pi} \left(\frac{m!}{\frac{m}{2}!} \right) C\Gamma \left(\frac{m-\alpha}{2} \right) (\lambda_+^{\alpha-m} + \lambda_-^{\alpha-m}) \text{ if } m = 2, 4, 6, \ldots$$

The mean, variance, skewness and excess kurtosis are given by

$$E(Z) = c_1(Z) = \mu + 2^{-\frac{\alpha+1}{2}} C\Gamma \left(\frac{1-\alpha}{2} \right) (\lambda_+^{\alpha-1} - \lambda_-^{\alpha-1}),$$

$$V(Z) = c_2(Z) = 2^{-\frac{\alpha+1}{2}} \sqrt{\pi} C\Gamma \left(1 - \frac{\alpha}{2} \right) (\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2}),$$

$$s(Z) = \frac{c_3(Z)}{c_2(Z)^{3/2}} = \frac{2^{\frac{\alpha+9}{4}} \Gamma \left(\frac{3-\alpha}{2} \right) (\lambda_+^{\alpha-3} - \lambda_-^{\alpha-3})}{\pi^{3/4} C^{1/2} (\Gamma(\frac{1-\alpha}{2})(\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2}))^{3/2}},$$

$$\kappa(Z) = \frac{c_4(Z)}{c_2(Z)^2} = \frac{3 \cdot 2^{\frac{\alpha+3}{2}} C\Gamma \left(2 - \frac{\alpha}{2} \right) (\lambda_+^{\alpha-4} + \lambda_-^{\alpha-4})}{\sqrt{\pi} C(\Gamma(\frac{1-\alpha}{2})(\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2}))^2}.$$

If $\alpha \in (0,2)\backslash\{1\}$, the Levy measure of α-stable, α-TS and α-MTS have the same asymptotic behavior at the zero neighborhood. However, the tails of the Levy measures for the α-MTS distribution are *thinner* than those of α-stable and *heavier* than those of α-TS distribution.

The moment estimators when Z is a Gamma process are given by

$$\hat{\theta}_n := \frac{\frac{1}{n^2} \left[\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta}) \right]^2}{\frac{1}{n^2} \sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})^2 - \frac{\Delta}{n} \left[\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta}) \right]} \frac{2a^3(a+1)}{b^4 \Delta},$$

$$\hat{\rho}_n := \frac{\frac{1}{n^2} \sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})^2 - \frac{\Delta}{n} \left[\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta}) \right]}{\frac{1}{n^2} \left[\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta}) \right]} \frac{b^3 \Delta}{2a^2(a+1)}.$$

For the MTS-OU model, the estimating functions are given by

$$c_1(y_1) = \lambda\rho\Delta c_1(Z), \quad c_2(y_1) = \Delta c_1(Z) + 2\lambda\rho^2 \Delta c_2(Z),$$

$$c_3(y_1) = \Delta c_1(Z) + 2\lambda\rho^2 \Delta c_2(Z), \quad c_4(y_1) = \Delta c_1(Z) + 2\lambda\rho^2 \Delta c_2(Z),$$

$$E(y_1) = c_1(y_1) = \mu + 2^{-\frac{\alpha+1}{2}} C\Gamma \left(\frac{1-\alpha}{2} \right) (\lambda_+^{\alpha-1} - \lambda_-^{\alpha-1}),$$

$$V(y_1) = c_2(y_1) = 2^{-\frac{\alpha+1}{2}} \sqrt{\pi} C\Gamma \left(1 - \frac{\alpha}{2} \right) (\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2}).$$

$$\hat{\theta}_n := \frac{\frac{1}{n^2}\left[\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})\right]^2}{\frac{1}{n^2}\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})^2 - \frac{\Delta}{n}\left[\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})\right]}$$

$$\times \ [2^{-\frac{\alpha+1}{2}} C\Gamma\left(\frac{1-\alpha}{2}\right)(\lambda_+^{\alpha-1} - \lambda_-^{\alpha-1})]^2 [2^{-\frac{\alpha+1}{2}}\sqrt{\pi} C\Gamma\left(1-\frac{\alpha}{2}\right)(\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2})] 2\Delta^{-1},$$

$$\hat{\rho}_n := \frac{\frac{1}{n^2}\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})^2 - \frac{\Delta}{n}\left[\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})\right]}{\frac{1}{n^2}\left[\sum_{i=1}^n (Y_{i\Delta} - Y_{(i-1)\Delta})\right]}$$

$$\times \ [2^{-\frac{\alpha+1}{2}} C\Gamma\left(\frac{1-\alpha}{2}\right)(\lambda_+^{\alpha-1} - \lambda_-^{\alpha-1}) 2^{-\frac{\alpha+1}{2}}\sqrt{\pi} C\Gamma\left(1-\frac{\alpha}{2}\right)(\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2})]^{-1} 2^{-1}\Delta.$$

Let $\vartheta = (\rho, \theta)$ and $\hat{\vartheta}_n = (\hat{\rho}_n, \hat{\theta}_n)$. By using Theorem 2.2 in Masuda [47] (see also Theorem 4.1 Van der Vaart [55]), we obtain the strong consistency and asymptotic normality of the MM estimators:

Proposition 5.1 For fixed $\Delta > 0$ as $n \to \infty$,

$$(a) \quad \hat{\vartheta}_n \to \vartheta_0 \ \text{ a.s.} \quad \text{as } n \to \infty,$$

$$(b) \quad \sqrt{n}(\hat{\vartheta}_n - \vartheta_0) \to^{\mathcal{D}} \mathcal{N}_2(0, (J^{-1}(\vartheta_0)) \text{ as } n \to \infty$$

where $J(\vartheta_0)$ Fisher information matrix.

6. Concluding Remarks

Financial return data are far from being Gaussian. We emphasized importance of non-Gaussian heavy tailed distributions in finance. We reviewed option pricing for stochastic interest rate and stochastic volatility models. We studied robust estimation in inverse Gaussian stochastic volatility model. We studied method of moments estimation in Gamma stochastic volatility model and modified tempered stable stochastic volatility model. The estimators are explicit and depend only on the asset price data. The estimation the other unknown parameters of the other processes based on asset price data in the hybrid model remains to be investigated which we would like to pursue in a future paper.

References

[1] Barndorff-Neilsen, O.E. and Shephard, N., 1998. Processes of normal inverse Gaussian type. Finance and Stochastics 2, 41-68.

[2] Barndorff-Nielsen, O.E. and Shephard, N., 2001. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics (with discussion). Journal of the Royal Statistical Society, Series B, 63, 167-241.

[3] Barndorff-Neilsen, O.E. and Shephard, N., 2002. Normal modified stable processes. Theory. Probab. Math. Statist. 65, 7-20.

[4] Ben Alaya, M. and Kebaier, A., 2012. Parameter estimation for the square-root diffusions: Ergodic and nonergodic cases. Stochastic Models 28, 609-634.

[5] Bender, C. and Marquardt, T., 2017. Stochastic calculus for convoluted Levy processes. Bernoulli 14, 499-518.

[6] Bishwal, J.P.N., 2006. Rates of weak convergence of the approximate minimum contrast estimators for the discretely observed Ornstein-Uhlenbeck process. Statistics and Probability Letters 76 (13), 1397-1409.

[7] Bishwal, J.P.N., 2007. A new estimating function for discretely sampled diffusions. Random Operators and Stochastic Equations 15 (1), 65-88

[8] Bishwal, J.P.N., 2008. Parameter Estimation in Stochastic Differential Equations, Lecture Notes in Mathematics 1923, Springer-Verlag.

[9] Bishwal, J.P.N., 2009. M-Estimation for discretely sampled diffusions Theory of Stochastic Processes 15 (31)(2), 62-83.

[10] Bishwal, J.P.N., 2010. Uniform rate of weak convergence for the minimum contrast estimator in the Ornstein-Uhlenbeck process. Methodology and Computing in Applied Probability 12 (3), 323-334.

[11] Bishwal, J.P.N., 2010. Sequential Monte Carlo methods for stochastic volatility models: a review. Journal of Interdisciplinary Mathematics 13 (6), 619-635.

[12] Bishwal, J.P.N., 2011. Sufficiency and Rao-Blackwellization of Vasicek model. Theory of Stochastic Processes 17 (1) , 12-15.

[13] Bishwal, J.P.N., 2011. Maximum quasi-likelihood estimation in fractional Levy stochastic volatility model. Journal of Mathematical Finance 1 (3), 12-15.

[14] Bishwal, J.P.N., 2011. Berry-Esseen inequalities for the discretely observed Ornstein-Uhlenbeck-Gamma process. Markov Processes and Related Fields 17(1), 119-150.

[15] Bishwal, J.P.N., 2011. Minimum contrast estimation in fractional Ornstein-Uhlenbeck process: continuous and discrete sampling. Fractional Calculus and Applied Analysis 14 (3), 375-410.

[16] Bishwal, J.P.N., 2011. Some new estimators of integrated volatility. Open Journal of Statistics 1 (2), 74-80.

[17] Bishwal, J.P.N., 2011. Financial extremes: a short review. Advances and Applications in Statistics 25 (1), 1-14.

[18] Bishwal, J.P.N., 2017. Hypothesis testing for fractional stochastic partial differential equations with applications to neurophysiology and finance. Asian Research Journal of Mathematics 4 (1), 1-24.

[19] Bishwal, J.P.N., 2021. A new algorithm for approximate maximum likelihood estimation in sub-fractional Chan-Karolyi-Longstaff-Sanders model. Asian Journal of Probability and Statistics 13 (3), 62-88.

[20] Bishwal, J.P.N., 2022. Parameter Estimation in Stochastic Volatility Models, Springer Nature, Cham.

[21] Black, F. and Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637-659.

[22] Blumenthal, R.M. and Getoor, R.K., 1961. Sample functions of stochastic processes with stationary independent increments. Journal of Mathematics and Mechanics 10 (3), 493-516.

[23] Brockwell, P.J., Davis, R.A. and Yang, Y., 2007. Estimation of non-negative Levy driven Ornstein-Uhlenbeck processes. Journal of Applied Probability 44, 977-989.

[24] Carr, P., Geman, H., Madan, D. and Yor, M., 2002. The fine structure of asset returns: An empirical investigation. Journal of Business 75 (2), 305-332.

[25] Carr, P. and Madan, D.B., 1999. Option valuation using fast Fourier transform. Journal

of Computational Finance 2 (4), 61-73.

[26] Cox, J.C., Ingersoll, J.E. and Ross, S.A., 1985. The theory of term structure of interest rates. Econometrica 53, 363-384.

[27] Davis, R.A. and Dunsmuir, W.T.M., 1997. Least absolute deviation estimation for regression with ARMA errors. Journal of Theoretical Probability 10, 481-497.

[28] Davis, R.A., Knight, K. and Liu, J., 1992. M-estimation for autoregressions with infinite variance. Stochastic Processes and their Applications 40, 145-180.

[29] Davis, R.A. and McCormick, W.P., 1989. Estimation for first order autoregressive processes with positive or bounded innovations. Stochastic Processes and their Applications 31, 237-250.

[30] Engle, R. and Russell, J., 1998. The autoregressive conditional duration model: a new model for irregularly spaced data. Econometrica 66, 1127-1162.

[31] Fama, E.F. and French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3-56.

[32] Fama, E.F. and French, K.R., 1995. Size and book-to-market factors in earnings and reutrns. Journal of Finance 50, 131-151.

[33] Fama, E.F. and French, K.R., 2015. A five-factor asset price model. Journal of Financial Economics 116, 1-22.

[34] Gaver, D.P. and Lewis, P.A.W., 1980. First order autoregressive gamma sequences and point processes. Advances in Applied Probability, 12, 727-745.

[35] Glasserman, P., 2004. Monte Carlo Methods in Financial Engineering, Springer, New York.

[36] Gourieroux, C. and Jasiak, J., 2006. Autoregressive gamma process. Journal of Forecasting 25 (2), 129-152.

[37] Heston, S.L., 1993. A closed-form solution for options with stochastic volatility. Review of Financial Studies 6 (2), 327-343.

[38] Hu, Y. and Long, H., 2009. Least squares estimation of Ornstein-Uhlenbeck process driven by α-stable motions. Stochastic Processes and their Applications 119, 2465-2480.

[39] Jongbloed, G., van der Meulen, F.H., van der Vaart, A.W., 2005. Nonparametric inference for Levy driven Ornstein-Uhlenbeck processes. Bernoulli 11 (5), 759-791.

[40] Kim, Y.S., Rachev, S.T., Chung, D.M., Bianichi, M.L., 2008. A modified tempered stable distribution with volatility clustering, In: J.O. Soares, J. Pina, M. Catalao-Lopes. New Developments in Financial Modelling, Cambridge Scholars Publishing, Newcastle upon Tyne, UK.

[41] Knight, K., 1998. Limit distributions for L_1 regression estimators under general conditions. Annals of Statistics 26, 755-770.

[42] Koenker, R., 2005. Quantile Regression, Cambridge University Press, Cambridge.

[43] Ling, S., 2005. Self-weighted least absolute deviation estimation for infinite variance autoregressive models. Journal of Royal Statistical Society, Series B, Statistical Methodology 67, 381-393.

[44] Neilsen, B. and Shephard, N., 2003. Likelihood analysis of a first order autoregressive model with exponential innovations. Journal of Time Series Analysis 24 (3), 337-344.

[45] Marquardt, T., 2006. Fractional Levy processes with applicati on to long memory moving average processes. Bernoulli 12 (6), 1009-1126.

[46] Marquardt, T., 2007. Multivariate fractionally integrated CARMA processes. Journal of Multivariate Analysis 98, 1705-1725.

[47] Masuda, H., 2005. Classical method of moments for partially and discretely observed ergodic models. Statistical Inference for Stochastic Processes 8, 25-50.

[48] Masuda, H., 2010. Asymptotic mixed normality in the self-weighted LAD estimation of Levy OU process. Electronic Journal of Statistics 4, 525-565.

[49] Masuda, H., 2019. Non-Gaussian quasi-likelihood estimation of SDE driven by locally stable Levy process. Stochastic Processes and their Applications 129, 1013-1059.

[50] Masuda, H. and Uehara, Y., 2017. Two-step estimation of ergodic Levy driven SDE. Statistical Inference for Stochastic Processes 20, 105-127.

[51] Neilsen, B. and Shephard, N., 2003. Likelihood analysis of a first order autoregressive model with exponential innovations. Journal of Time Series Analysis 24 (3), 337-344.

[52] Portnoy, S. and Koenker, R., 1997. The Gaussian hare and Laplacian tortoise: computability of squared-error versus absolute-error estimators. Statistical Science 12, 279-300.

[53] Sharpe, W.F., 1964. Capital asset prices: A theory of Market equilibrium under conditions of risk. Journal of Finance, 19, 425-442.

[54] Sim, C., 1990. First order autoregressive model for gamma and exponential processes. Journal of Applied Probability, 27, 325-332.

[55] Van der Vaart, A.W., 2000. Asymptotic Statistics, Cambridge University Press, Cambridge.

[56] Wolfe, A.J., 1982. On a continuous analogue of the stochastic difference equation $X_n = \rho X_{n-1} + B_n$. Stochastic Processes and their Applications 12, 301-312.