North American Journal of Economics and Finance 56 (2021) 101326

Contents lists available at ScienceDirect

North American Journal of Economics and Finance

journal homepage: www.elsevier.com/locate/najef

Consistent pricing of VIX options with the Hawkes jump-diffusion model

Bo Jing [a], Shenghong Li [a], Yong Ma [b,*]

[a] School of Mathematical Sciences, Zhejiang University, Hangzhou 310027, PR China
[b] College of Finance and Statistics, Hunan University, Changsha, 410006, PR China

ARTICLE INFO

Keywords:
Consistent approach
VIX options
Hawkes process
Jump-diffusion
COS method

ABSTRACT

This paper presents a valuation of VIX options employing a Hawkes jump-diffusion model that captures the clustering pattern of jumps observed extensively in the financial markets. In the consistent framework, the valuation problem of VIX options is solved efficiently via the Fourier cosine expansion (COS) method. The Monte Carlo (MC) simulations are carried out to demonstrate the reliability and efficiency of the COS method. Furthermore, a sensitivity analysis is performed to show how option prices response to different parameters associated with jump clustering. Finally, empirical studies are conducted to provide evidence to support our jump specification in matching the VIX option surface.

1. Introduction

The volatility index (VIX), introduced by the Chicago Board Options Exchange (CBOE) in 1993, represents the expected market volatility implied by the options on the S&P 500 index (SPX) over the next 30 calendar days. As the VIX itself cannot be traded directly, VIX derivatives, like VIX futures and option, as the tradable assets provide investors with new instruments for trading volatility directly. Since the CBOE launched VIX options in 2006, the popularity of trading these products has been growing. The main reason is that market practitioners can gain exposure to the volatility of the SPX and hedge against the risks of potential market downturns through the trading of VIX options. Nowadays, VIX options are among the most actively traded products on CBOE, with the average daily trading volume of approximately 666,161 contracts in 2018 (see Fig. 1). With the rapid growth of the VIX derivatives market, developing a reliable pricing model has become an important topic in both empirical and theoretical research. However, pricing VIX options is a challenging endeavor. This new asset class has different characteristics from stock indexes, such as mean reversion, so that the option pricing models originally developed for stock indexes is no longer applicable. To date, abundant literature has emerged on the valuation of VIX derivatives. According to different modeling ideas, these studies are classified into two main categories: a standalone and a consistent approach. The first category directly models the VIX dynamics as a stochastic process (e.g., Goard & Mazur, 2013; Grünbichler & Longstaff, 1996; Mencía & Sentana, 2013; Park, 2016; Psychoyios et al., 2010). The second one starts from specifying the dynamics of the underlying SPX, and then the VIX can be derived from the underlying (e.g., Bardgett et al., 2019; Luo et al., 2019; Pacati et al., 2018; Zhang & Zhu, 2006; Zhu & Lian, 2012). In this research stream, the challenge lies in ensuring that empirical characteristics of the joint dynamics of SPX and VIX are captured. In this paper, we adopt the consistent approach to solve the pricing problem for VIX options and consider the clustering pattern of jumps by employing a Hawkes jump-diffusion model. On the other hand, preserving an acceptable level of tractability is also a problem that needs to be addressed when pricing derivatives. The model specification makes it challenging to implement and we successfully address this problem by developing the efficient COS method.

* Corresponding author.
E-mail addresses: bojing@zju.edu.cn (B. Jing), shli@zju.edu.cn (S. Li), yma@hnu.edu.cn (Y. Ma).

https://doi.org/10.1016/j.najef.2020.101326
Received 17 January 2020; Received in revised form 19 August 2020; Accepted 16 November 2020
Available online 20 November 2020

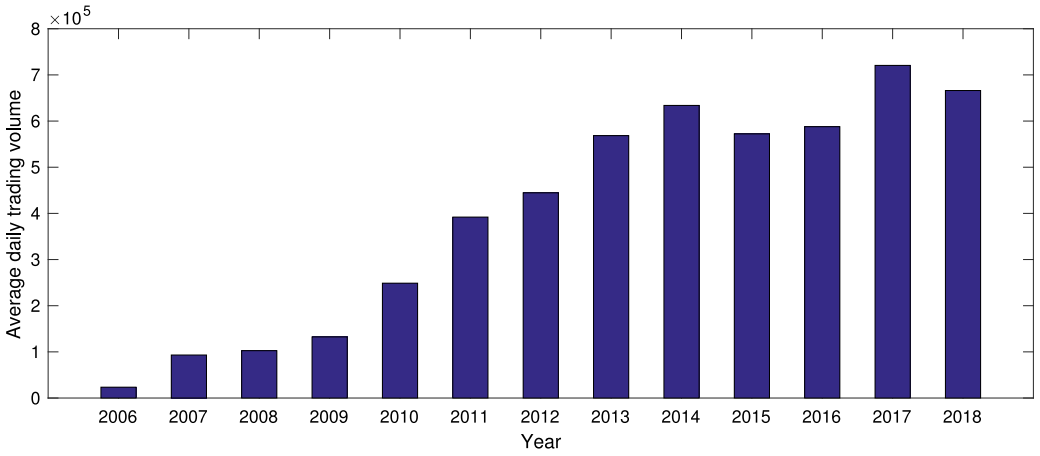


Fig. 1. Average daily trading volume of VIX options by year.

It is widely acknowledged that both stochastic volatility and random jumps play important roles in derivatives pricing. Yet traditional jump-diffusion models are inconsistent with empirical observations that extreme and sudden price movements, referred to as jumps, tend to concentrate in short spreads of time, generating the clustering pattern of jumps. Recent literature documents that the occurrence of one jump increases the likelihood of future jumps, causing jumps to occur in clusters, which appears even more pronounced in time of financial turmoil. For example, Fulop et al. (2015) found that the jumps clustering phenomenon is particularly evident during the market crash of 1987 and the subprime mortgage crisis of 2008. Du and Luo (2019) provided strong evidence for the existence of jump clustering uncovered from the SPX returns. To capture the empirical pattern, one of the most popular modeling methods is the use of the Hawkes process after Hawkes (1971), which is ideal for dealing with the clustering of events owing to its self-exciting property. The Hawkes process is originally used for the modeling of earthquake occurrences, which are highly clustered in time since a quake is likely to be followed by aftershocks (see Ogata, 1988).

In recent years, the Hawkes process has gained popularity in the financial field (e.g., Bacry & Muzy, 2014; Chavez-Demoulin & McGill, 2012; Errais et al., 2010; Herrera & Schipp, 2014). One of the most influential pioneer works is Aït-Sahalia et al. (2015), which first introduced Hawkes processes to model international equity index returns and developed a Hawkes jump-diffusion model via adding a continuous Brownian component with stochastic volatility. Since then, the applications of Hawkes processes are especially successful in the domain of the research on asset pricing (e.g., Kokholm, 2016; Liu & Zhu, 2019; Ma et al., 2017; Zhang et al., 2017). As stated by Liu and Zhu (2019), the Lévy processes, even having very general form, are incapable of reproducing clustered jumps because of their independent increment property. However, the Hawkes jump-diffusion model combines the observed market features of stochastic volatility and jump clustering via a self-exciting Hawkes process. Therefore, inspired by these previous studies, this paper employs the Hawkes jump-diffusion model in the context of VIX option pricing to accommodate the empirically observed market features. To the best of our knowledge, there are no studies on pricing VIX options with the Hawkes jump-diffusion model in the consistent framework. This paper aims to fill this vacuum in the finance literature.

The main contributions of this paper are three-fold. First, this paper adds to the existing literature by applying the Hawkes jump-diffusion model to VIX option pricing in the consistent framework. The empirically observed pattern of jump clustering, which has been previously ignored in much of the VIX option pricing literature, can therefore be captured by a self-exciting Hawkes process. Second, the valuation problem of VIX options is solved via the COS method, which is an efficient pricing method introduced by Fang and Oosterlee (2008). This paper extends the COS method to VIX options pricing with self-exciting jumps. Finally, numerical experiments are conducted to examine the reliability and efficiency of the pricing method and explore the impact of jump clustering on option prices. Then, to illustrate the ability of our jump specification to match the VIX option surface, empirical studies are performed in which the results are compared with the benchmark model in the aspect of fitting the market data. Empirical results suggest that our jump specification inducing jump clustering indeed has a significant impact on VIX option prices, since the Hawkes jump-diffusion model provides a better fit to the option price surface than the benchmark model.

The rest of this paper is organized as follows. Section 2 introduces the model setup and some facts of the CBOE VIX. The characteristic function is derived in Section 3. Section 4 details the COS method for pricing VIX options. In Section 5, we exhibit numerical experiments to demonstrate the reliability and efficiency of the COS method and give a sensitivity analysis of option prices with respect to different parameters associated with jump clustering. Section 6 describes the market data used for empirical studies and reports the results of the model comparisons. The last section concludes.

2. The model

2.1. Model setup

Let $(\Omega, \mathcal{F}, \mathbb{Q})$ be a complete probability space, equipped with an information flow $\{\mathcal{F}_t\}_{t\geq 0}$, where $\mathbb{Q}$ is a risk-neutral probability measure. On the filtered probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, the dynamics of the SPX, denoted by S_t, follow the Hawkes jump-diffusion

model:

$$\begin{aligned} \frac{dS_t}{S_{t-}} &= (r-q-\bar{\mu}\lambda_t)dt + \sqrt{V_t}dW_t^S + (e^{Y_t}-1)dN_t, \\ dV_t &= \kappa\left(\theta - V_t\right)dt + \sigma\sqrt{V_t}dW_t^V, \\ d\lambda_t &= \delta\left(\bar{\lambda}-\lambda_t\right)dt + \epsilon dN_t, \end{aligned} \tag{1}$$

where W_t^S and W_t^V are a pair of correlated standard Brownian motions with correlation coefficient ρ; r is the risk-free interest rate; q is the continuous dividend yield; κ, θ and σ represent the mean-reversion speed, the long-run mean level and the volatility of the instantaneous variance V_t, respectively; $e^{Y_t}-1$ is the percentage price jump size with mean $\bar{\mu}$. We assume that Y_t is a family of independent and identically distributed random variables with $Y_t \sim \mathcal{N}(\mu_t, \sigma_t^2)$ for any fixed t, and then $\bar{\mu} = E(e^{Y_t}-1) = e^{\mu_s+\frac{1}{2}\sigma_s^2}-1$; N_t is the Hawkes process with stochastic intensity λ_t. The characteristic feature of the Hawkes process is its self-exciting property: the occurrence of one jump would increase the jump intensity by ϵ and then the intensity decays exponentially over time at a speed δ to the background intensity $\bar{\lambda}$, which is a simple but popular way to capture the clustering pattern of jumps; Finally, we assume that (W_t^S, W_t^V), N_t and Y_t are mutually independent. We refer to the model as the SV-HJ model hereafter.

2.2. The CBOE VIX

The VIX was originally calculated based on the implied volatilities of eight at-the-money S&P 100 index options with maturities closest to 30 calendar days. On September 23, 2003, the CBOE together with Goldman Sachs updated the methodology for calculating the VIX and the index using the old methodology was renamed the VXO. As described in the CBOE white paper, the VIX is now defined by:

$$\left(\frac{VIX_t}{100}\right)^2 = \frac{2}{\bar{\tau}}\sum_i \frac{\Delta K_i}{K_i^2}e^{r\bar{\tau}}Q_t(K_i) - \frac{1}{\bar{\tau}}\left(\frac{F_t}{K_0}-1\right)^2, \tag{2}$$

where $\bar{\tau} = \frac{30}{365}$, K_i is the strike price of the ith out-of-the-money SPX option, $Q_t(K_i)$ is the mid-quote of the option with strike K_i at time t, K_0 is the highest strike price below the time-t forward price F_t, r is the time-t risk-free interest rate, and ΔK_i is the strike price increment given by $\Delta K_i = \frac{K_{i+1}-K_{i-1}}{2}$. It can be seen that the current VIX is calculated in a model-free manner using a portfolio of call and put options on the SPX. The portfolio weights match those of the 30-day log contract replication. Thus, the squared VIX can be expressed, with mathematical simplification, in terms of the risk-neutral expectation of the log contract (see Duan & Yeh, 2010; Lin, 2007; Zhu & Zhang, 2007 for more details):

$$\left(\frac{VIX_t}{100}\right)^2 = -\frac{2}{\bar{\tau}}E_t^{\mathbb{Q}}\left[\log\left(\frac{S_{t+\bar{\tau}}}{F_t}\right)\right], \tag{3}$$

where $F_t = S_t e^{(r-q)\bar{\tau}}$ is the time-t forward price of the SPX with maturity $t+\bar{\tau}$, and $E_t^{\mathbb{Q}}[\cdot]$ denotes the expectation conditional on the information available up to time t under the risk neutral measure $\mathbb{Q}$. For notational convenience, let VIX_t' be the VIX expressed in percentage points, i.e., $VIX_t' = VIX_t/100$. The following proposition shows that $VIX_t'^2$ has the analytic representation in terms of V_t and λ_t.

Proposition 2.1. *Under the SV-HJ model described in (1), the squared VIX at time t can be expressed as an affine function of V_t and λ_t:*

$$VIX_t'^2 = \alpha V_t + \beta\lambda_t + \gamma, \tag{4}$$

where the coefficients α, β and γ are given by:

$$\begin{aligned} \alpha &= \frac{1-e^{-\kappa\bar{\tau}}}{\kappa\bar{\tau}}, \\ \beta &= 2(\bar{\mu}-\mu_s)\frac{1-e^{-(\delta-\epsilon)\bar{\tau}}}{(\delta-\epsilon)\bar{\tau}}, \\ \gamma &= \theta(1-\alpha) + \frac{\delta\bar{\lambda}}{\delta-\epsilon}[2(\bar{\mu}-\mu_s)-\beta]. \end{aligned}$$

Proof. For any $0 \le t < s$, applying Itô's lemma to calculate $d(e^{-\kappa(s-u)}V_u)$ with respect to u and integrating both sides from t to s, we have:

$$V_s = e^{-\kappa(s-t)}V_t + (1-e^{-\kappa(s-t)})\theta + \sigma\int_t^s e^{-\kappa(s-u)}\sqrt{V_u}dW_u^V.$$

Taking the conditional expectation on both sides of the above equation, we get the first moment of V_s as:

$$E_t^{\mathbb{Q}}(V_s) = e^{-\kappa(s-t)}V_t + (1-e^{-\kappa(s-t)})\theta.$$

The dynamics of λ_t can be rewritten as:

$$d\lambda_t = \kappa_\lambda\left(\theta_\lambda - \lambda_t\right)dt + \epsilon dM_t,$$

where $M_t = N_t - \int_0^t \lambda_u du$ is a local martingale, $\kappa_\lambda = \delta - \epsilon$ and $\theta_\lambda = \frac{\delta\bar{\lambda}}{\delta-\epsilon}$. For any $0 \le t < s$, applying Itô's lemma to $e^{-\kappa_\lambda(s-u)}\lambda_u$ yields:

$$\lambda_s = e^{-\kappa_\lambda(s-t)}\lambda_t + (1 - e^{-\kappa_\lambda(s-t)})\theta_\lambda + \epsilon \int_t^s e^{-\kappa_\lambda(s-u)} dM_u.$$

Taking the conditional expectation on both sides of the above equation, we get the first moment of λ_s as:

$$E_t^{\mathbb{Q}}(\lambda_s) = e^{-\kappa_\lambda(s-t)}\lambda_t + (1 - e^{-\kappa_\lambda(s-t)})\theta_\lambda.$$

By Itô's lemma, the dynamics of the log-price with respect to M_t can be expressed as:

$$d\log(S_t) = \Big(r - q - \frac{1}{2}V_t - (\bar{\mu} - Y_t)\lambda_t\Big)dt + \sqrt{V_t}dW_t^S + Y_t dM_t.$$

Integrating both sides from t to $t + \bar{\tau}$, we have:

$$\log(\frac{S_{t+\bar{\tau}}}{S_t}) = \int_t^{t+\bar{\tau}} \Big(r - q - \frac{1}{2}V_s - (\bar{\mu} - Y_s)\lambda_s\Big)ds + \int_t^{t+\bar{\tau}} \sqrt{V_s}dW_s^S + \int_t^{t+\bar{\tau}} Y_s dM_s.$$

Taking the conditional expectation on both sides of the above equation and then using the first moment of V_s and λ_s yields:

$$E_t^{\mathbb{Q}}\left[\log\Big(\frac{S_{t+\bar{\tau}}}{S_t}\Big)\right] = [r - q - \frac{1}{2}\theta - (\bar{\mu} - \mu_s)\theta_\lambda]\bar{\tau} - \frac{V_t - \theta}{2\kappa}(1 - e^{-\kappa\bar{\tau}}) - \frac{(\bar{\mu} - \mu_s)(\lambda_t - \theta_\lambda)}{\kappa_\lambda}(1 - e^{-\kappa_\lambda\bar{\tau}}).$$

According to its definition in (3), the squared VIX can be expressed as:

$$VIX_t'^2 = \alpha V_t + \beta\lambda_t + \gamma,$$

where the coefficients α, β and γ are given by:

$$\alpha = \frac{1 - e^{-\kappa\bar{\tau}}}{\kappa\bar{\tau}},$$
$$\beta = 2(\bar{\mu} - \mu_s)\frac{1 - e^{-\kappa_\lambda\bar{\tau}}}{\kappa_\lambda\bar{\tau}},$$
$$\gamma = \theta(1 - \alpha) + \theta_\lambda[2(\bar{\mu} - \mu_s) - \beta].$$

Substituting $\kappa_\lambda = \delta - \epsilon$ and $\theta_\lambda = \frac{\delta\bar{\lambda}}{\delta-\epsilon}$ into the above equation, we obtain the desired result. □

It is noted that some classic pricing models can be considered as special cases of the general model presented here. If we assume no jump risk and obviously no jump clustering (i.e., $\mu_s = \sigma_s = \lambda_t = 0$), then the general model degenerates to the Heston model, which is a successful stochastic volatility model in the literature due to its analytical tractability. It can be shown that the formula derived in Zhang and Zhu (2006) is indeed the same as the above one. If there is no jump clustering, that is, N_t is a Poisson process with constant intensity $\lambda_t = \lambda$, the general model leads to the Poisson jump-diffusion model with stochastic volatility. In this case, the above formula becomes $VIX_t'^2 = \alpha V_t + \beta$ with $\alpha = \frac{1-e^{-\kappa\bar{\tau}}}{\kappa\bar{\tau}}$ and $\beta = \theta(1 - \alpha) + 2(\bar{\mu} - \mu_s)\lambda$, which is a simplified form of the corresponding results presented in Zhu and Lian (2012). They also incorporate the variance jumps into model. Here we restrict to a parsimonious specification and leave the extension to include the variance jumps for future research.

3. The characteristic function

Let us mention that the SV-HJ model falls into the class of affine jump-diffusion models. Therefore, we apply the technique of Duffie et al. (2000) to derive the characteristic function. The time-t conditional characteristic function of $VIX_T'^2$ under the risk-neutral measure $\mathbb{Q}$ is defined as:

$$\Phi\left(\phi; V_t, \lambda_t, \tau\right) = E_t^{\mathbb{Q}}\left(e^{i\phi VIX_T'^2}\right), \tag{5}$$

where $\tau = T - t$ and $i = \sqrt{-1}$. The characteristic function can be obtained by solving a system of ordinary differential equations, as presented in the following proposition.

Proposition 3.1. *Under the SV-HJ model described in (1), the time-t conditional characteristic function of $VIX_T'^2$ is given by:*

$$\Phi\left(\phi; V_t, \lambda_t, \tau\right) = e^{A(\phi;\tau)V_t + B(\phi;\tau)\lambda_t + C(\phi;\tau)}, \tag{6}$$

where $B(\phi;\tau)$ and $C(\phi;\tau)$ satisfy the following system of ordinary differential equations (ODEs):

$$\frac{\partial B(\phi;\tau)}{\partial\tau} = -\delta B(\phi;\tau) + e^{\epsilon B(\phi;\tau)} - 1,$$
$$\frac{\partial C(\phi;\tau)}{\partial\tau} = \kappa\theta A(\phi;\tau) + \delta\bar{\lambda}B(\phi;\tau),$$

subject to the boundary conditions $B(\phi;0)=i\phi\beta$ *and* $C(\phi;0)=i\phi\gamma$, *and* $A(\phi;\tau)$ *has the analytic expression:*

$$A(\phi;\tau)=\frac{2\kappa\alpha\phi i}{\sigma^2\alpha\phi i+(2\kappa-\sigma^2\alpha\phi i)e^{\kappa\tau}}.$$

Proof. Applying the Feynman–Kac theorem, $\Phi(\cdot)$ is governed by the following partial integro-differential equation (PIDE):

$$\begin{aligned}\frac{\partial\Phi}{\partial\tau}=&\kappa(\theta-V_t)\frac{\partial\Phi}{\partial V_t}+\delta(\bar{\lambda}-\lambda_t)\frac{\partial\Phi}{\partial\lambda_t}+\frac{1}{2}\sigma^2V_t\frac{\partial^2\Phi}{\partial V_t^2}\\&+\lambda_tE[\Phi(\phi;V_t,\lambda_t+\epsilon,\tau)-\Phi(\phi;V_t,\lambda_t,\tau)],\end{aligned}$$

with the terminal condition $\Phi(\phi;V_T,\lambda_T,0)=e^{i\phi VIX_T'^2}$. Using Proposition 2.1, we have:

$$\Phi(\phi;V_T,\lambda_T,0)=e^{i\phi\alpha V_T+i\phi\beta\lambda_T+i\phi\gamma}.$$

Borrowing the idea of Duffie et al. (2000), we conjecture that $\Phi(\cdot)$ has the exponential affine form as follows:

$$\Phi\left(\phi;V_t,\lambda_t,\tau\right)=e^{A(\phi;\tau)V_t+B(\phi;\tau)\lambda_t+C(\phi;\tau)}.$$

Substituting this solution form into the above PIDE yields:

$$\begin{aligned}\frac{\partial A(\phi;\tau)}{\partial\tau}V_t+\frac{\partial B(\phi;\tau)}{\partial\tau}\lambda_t+\frac{\partial C(\phi;\tau)}{\partial\tau}=&\kappa(\theta-V_t)A(\phi;\tau)+\delta(\bar{\lambda}-\lambda_t)B(\phi;\tau)\\&+\frac{1}{2}\sigma^2V_tA(\phi;\tau)^2+\lambda_t(e^{\epsilon B(\phi;\tau)}-1).\end{aligned}$$

Regrouping the equation with respect to V_t and λ_t, we have:

$$\begin{aligned}0=&\left(-\kappa A(\phi;\tau)+\frac{1}{2}\sigma^2A(\phi;\tau)^2-\frac{\partial A(\phi;\tau)}{\partial\tau}\right)V_t\\&+\left(-\delta B(\phi;\tau)+e^{\epsilon B(\phi;\tau)}-1-\frac{\partial B(\phi;\tau)}{\partial\tau}\right)\lambda_t\\&+\kappa\theta A(\phi;\tau)+\delta\bar{\lambda}B(\phi;\tau)-\frac{\partial C(\phi;\tau)}{\partial\tau}.\end{aligned}$$

Since V_t and λ_t are stochastic, to make the right-hand side of the above equation always equal to zero, we obtain the following system of ordinary differential equations (ODEs):

$$\begin{aligned}\frac{\partial A(\phi;\tau)}{\partial\tau}&=-\kappa A(\phi;\tau)+\frac{1}{2}\sigma^2A(\phi;\tau)^2,\\\frac{\partial B(\phi;\tau)}{\partial\tau}&=-\delta B(\phi;\tau)+e^{\epsilon B(\phi;\tau)}-1,\\\frac{\partial C(\phi;\tau)}{\partial\tau}&=\kappa\theta A(\phi;\tau)+\delta\bar{\lambda}B(\phi;\tau),\end{aligned}$$

subject to the boundary conditions $A(\phi;0)=i\phi\alpha$, $B(\phi;0)=i\phi\beta$ and $C(\phi;0)=i\phi\gamma$.

The first ODE known as the Riccati equation has analytical solution as follows:

$$A(\phi;\tau)=\frac{2\kappa\alpha\phi i}{\sigma^2\alpha\phi i+(2\kappa-\sigma^2\alpha\phi i)e^{\kappa\tau}},$$

In general, there are no closed-form solutions for $B(\phi;\tau)$ and $C(\phi;\tau)$. However, if we reduce the Hawkes processes to Poisson processes, then the closed-form solutions are available, which corresponds to those of Zhu and Lian (2012). In this paper, we will solve the ODEs numerically using the Runge–Kutta algorithm. □

4. Pricing VIX options using the COS method

Once the characteristic function is available, the problem of option pricing can be usually performed using Fourier inversion techniques. However, it can be seen from the representation of the VIX that the logarithm of the VIX is not affine, which makes it impossible to apply the fast Fourier transform (FFT) method of Carr and Madan (1999). To solve the valuation problem of VIX options, we resort to the COS method, which was originally introduced by Fang and Oosterlee (2008) and is a fast and efficient pricing method with exponential convergence and linear computational complexity. The COS method has been successfully applied in valuing financial derivatives (e.g., Chau et al., 2015; Li et al., 2019; Tour et al., 2018; Zhang & Oosterlee, 2014).

We consider a VIX call option with strike price K and maturity T, whose terminal payoff is described as $(VIX_T-K)^+$. According to the risk-neutral valuation theory, the option price at time t, denoted by C_t, can be expressed as the discounted conditional expectation of the terminal payoff under the $\mathbb{Q}$:

$$\begin{aligned}C_t&=e^{-r\tau}E_t^{\mathbb{Q}}\left[\left(VIX_T-K\right)^+\right]\\&=100e^{-r\tau}E_t^{\mathbb{Q}}\left[\left(\sqrt{VIX_T'^2}-K'\right)^+\right]\\&=100e^{-r\tau}\int_0^{+\infty}\left(\sqrt{y}-K'\right)^+f_\tau(y|V_t,\lambda_t)\,dy,\end{aligned}\tag{7}$$

where $\tau = T - t$, K' is the strike price expressed in percentage points and $f_\tau(y|V_t, \lambda_t)$ is the risk-neutral conditional probability density function of $VIX_T'^2$ given the information available up to time t. By the inverse Fourier transform, the conditional density function can be given by:

$$f_\tau(y|V_t, \lambda_t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\phi y} \Phi(\phi; V_t, \lambda_t, \tau) d\phi. \tag{8}$$

Given the suitable choice of the truncation range, which will be detailed later, $f_\tau(y|V_t, \lambda_t)$ can be approximated by a truncated Fourier cosine series expansion:

$$f_\tau(y|V_t, \lambda_t) \approx \frac{2}{b-a} {\sum_{n=0}^{N-1}}' \Re \left[\Phi\left(\frac{n\pi}{b-a}; V_t, \lambda_t, \tau\right) e^{-i\frac{na\pi}{b-a}} \right] cos\left(n\pi \frac{y-a}{b-a}\right), \tag{9}$$

where $\sum'$ means that the first term in the summation is multiplied by one-half and $\Re[\cdot]$ denotes the real part of the argument. Substituting (9) into (7) and interchanging the integration and summation gives the pricing formula for the VIX option:

$$C_t \approx 100 e^{-r\tau} {\sum_{n=0}^{N-1}}' \Re \left[\Phi\left(\frac{n\pi}{b-a}; V_t, \lambda_t, \tau\right) e^{-i\frac{na\pi}{b-a}} \right] U_n, \tag{10}$$

where

$$U_n = \frac{2}{b-a} \int_a^b \left(\sqrt{y} - K'\right)^+ cos\left(n\pi \frac{y-a}{b-a}\right) dy. \tag{11}$$

The following lemma shows that the payoff series coefficients, U_n, can be obtained analytically for VIX options.

Lemma 4.1. *The Fourier cosine series coefficients, U_n, of payoff $\left(\sqrt{y} - K'\right)^+$ on $[a, b]$ are known analytically:*

For $n = 0$,

$$U_n = \begin{cases} \frac{2}{b-a}\left(\frac{2}{3}b^{\frac{3}{2}} - K'b + \frac{1}{3}K'^3\right), & a < K'^2, \\ \frac{2}{b-a}\left(\frac{2}{3}b^{\frac{3}{2}} - \frac{2}{3}a^{\frac{3}{2}} - K'b + K'a\right), & a \geq K'^2; \end{cases}$$

For $n \neq 0$,

$$U_n = \begin{cases} \frac{2}{b-a}\Re\left\{ e^{-i\omega_n a}\left[\frac{\sqrt{b}-K'}{i\omega_n} e^{i\omega_n b} + \frac{\sqrt{\pi}}{2(\sqrt{-i\omega_n})^3} \right.\right. \\ \left.\left. \cdot \left(erfz(\sqrt{-i\omega_n b}) - erfz(K'\sqrt{-i\omega_n}) \right) \right] \right\}, & a < K'^2, \\ \frac{2}{b-a}\Re\left\{ e^{-i\omega_n a}\left[\frac{\sqrt{b}-K'}{i\omega_n} e^{i\omega_n b} - \frac{\sqrt{a}-K'}{i\omega_n} e^{i\omega_n a} + \frac{\sqrt{\pi}}{2(\sqrt{-i\omega_n})^3} \right.\right. \\ \left.\left. \cdot \left(erfz(\sqrt{-i\omega_n b}) - erfz(\sqrt{-i\omega_n a}) \right) \right] \right\}, & a \geq K'^2, \end{cases}$$

where $\omega_n = \frac{n\pi}{b-a}$ and $erfz(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$ is the complex error function.

Proof. For notational convenience, we denote $\omega_n = \frac{n\pi}{b-a}$. For $n \neq 0$, if $a < K'^2$, then we have:

$$\begin{aligned} U_n &= \frac{2}{b-a} \int_{K'^2}^b \left(\sqrt{y} - K'\right) cos\left(\omega_n (y-a)\right) dy \\ &= \frac{2}{b-a} \Re\left[\int_{K'^2}^b \sqrt{y} e^{i\omega_n (y-a)} dy - K' \int_{K'^2}^b e^{i\omega_n (y-a)} dy \right]. \end{aligned}$$

The first term in the above square bracket can be calculated as

$$\begin{aligned} \int_{K'^2}^b \sqrt{y} e^{i\omega_n (y-a)} dy &= e^{-i\omega_n a}\left[\frac{\sqrt{b}}{i\omega_n} e^{i\omega_n b} - \frac{K'}{i\omega_n} e^{i\omega_n K'^2} - \int_{K'}^{\sqrt{b}} \frac{1}{i\omega_n} e^{i\omega_n y^2} dy \right] \\ &= e^{-i\omega_n a}\left[\frac{\sqrt{b}}{i\omega_n} e^{i\omega_n b} - \frac{K'}{i\omega_n} e^{i\omega_n K'^2} + \frac{\sqrt{\pi}}{2(\sqrt{-i\omega_n})^3} \right. \\ &\quad \left. \cdot \left(erfz(\sqrt{-i\omega_n b}) - erfz(K'\sqrt{-i\omega_n}) \right) \right], \end{aligned}$$

where $erfz(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$ is the complex error function. The second term in the square bracket can be calculated as

$$K' \int_{K'^2}^b e^{i\omega_n (y-a)} dy = K' e^{-i\omega_n a}\left(\frac{1}{i\omega_n} e^{i\omega_n b} - \frac{1}{i\omega_n} e^{i\omega_n K'^2} \right).$$

Therefore, the payoff series coefficients U_n becomes:

$$U_n = \frac{2}{b-a}\Re\left\{ e^{-i\omega_n a}\left[\frac{\sqrt{b}-K'}{i\omega_n}e^{i\omega_n b} + \frac{\sqrt{\pi}}{2(\sqrt{-i\omega_n})^3} \cdot \left(erfz(\sqrt{-i\omega_n}b) - erfz(K'\sqrt{-i\omega_n})\right)\right]\right\}.$$

If $a \geq K'^2$, then we have:

$$U_n = \frac{2}{b-a}\int_a^b \left(\sqrt{y}-K'\right)cos\left(\omega_n(y-a)\right)dy.$$

The coefficients U_n can be obtained in an analogous manner as follows:

$$U_n = \frac{2}{b-a}\Re\left\{ e^{-i\omega_n a}\left[\frac{\sqrt{b}-K'}{i\omega_n}e^{i\omega_n b} - \frac{\sqrt{a}-K'}{i\omega_n}e^{i\omega_n a} + \frac{\sqrt{\pi}}{2(\sqrt{-i\omega_n})^3} \times \left(erfz(\sqrt{-i\omega_n}b) - erfz(\sqrt{-i\omega_n}a)\right)\right]\right\}.$$

For $n = 0$, it is straightforward to calculate the integral

$$U_n = \frac{2}{b-a}\int_a^b \left(\sqrt{y}-K'\right)^+ dy = \begin{cases} \frac{2}{b-a}\left(\frac{2}{3}b^{\frac{3}{2}} - K'b + \frac{1}{3}K'^3\right), & a < K'^2, \\ \frac{2}{b-a}\left(\frac{2}{3}b^{\frac{3}{2}} - \frac{2}{3}a^{\frac{3}{2}} - K'b + K'a\right), & a \geq K'^2. \end{cases} \quad \square$$

Following the choice of the truncation range $[a, b]$ proposed by Fang and Oosterlee (2008), the sole difference is that the lower bound, a, of the truncation range is chosen to be the maximum of $c_1 - L\sqrt{c_2 + \sqrt{c_4}}$ and 0, because the squared VIX should not be negative. The truncation range $[a, b]$ is thus determined as follows:

$$[a,b] = \left[\max(c_1 - L\sqrt{c_2+\sqrt{c_4}}, 0), c_1 + L\sqrt{c_2+\sqrt{c_4}}\right] \quad with\ L = 10. \tag{12}$$

Here c_j is the jth cumulant of $VIX_T'^2$ defined as:

$$c_j = \frac{1}{i^j}\frac{\partial^j(\log \Phi(\phi))}{\partial \phi^j}\bigg|_{\phi=0}, \tag{13}$$

where $\Phi(\cdot)$ is the characteristic function of $VIX_T'^2$. The analytic expressions of c_j, which are needed to determine the truncation range, are presented in the following lemma.

Lemma 4.2. *The first, second and fourth cumulants of $VIX_T'^2$ can be derived as:*

$$\begin{aligned}
c_1 &= \alpha e^{-\kappa\tau}V_t + \beta e^{-(\delta-\epsilon)\tau}\lambda_t + \alpha\theta(1-e^{-\kappa\tau}) + \frac{\beta\delta\bar{\lambda}}{\delta-\epsilon}(1-e^{-(\delta-\epsilon)\tau}) + \gamma,\\
c_2 &= \frac{\alpha^2\sigma^2}{\kappa}(e^{-\kappa\tau}-e^{-2\kappa\tau})V_t + \frac{\beta^2\epsilon^2}{\delta-\epsilon}(e^{-(\delta-\epsilon)\tau}-e^{-2(\delta-\epsilon)\tau})\lambda_t\\
&\quad + \frac{\alpha^2\theta\sigma^2}{2\kappa}(1-e^{-\kappa\tau})^2 + \frac{\beta^2\delta\bar{\lambda}\epsilon^2}{2(\delta-\epsilon)^2}(1-e^{-(\delta-\epsilon)\tau})^2,\\
c_4 &= \frac{3\alpha^4\sigma^6}{\kappa^3}e^{-\kappa\tau}(1-e^{-\kappa\tau})^3V_t + \left[\frac{3\beta^4\epsilon^6}{(\delta-\epsilon)^3}e^{-(\delta-\epsilon)\tau}(1-e^{-(\delta-\epsilon)\tau})^3\right.\\
&\quad + \frac{\beta^4\epsilon^5}{3(\delta-\epsilon)^2}\left(7e^{-(\delta-\epsilon)\tau} - 6e^{-2(\delta-\epsilon)\tau} - 9e^{-3(\delta-\epsilon)\tau} + 8e^{-4(\delta-\epsilon)\tau}\right)\\
&\quad \left. + \frac{\beta^4\epsilon^4}{3(\delta-\epsilon)}(e^{-(\delta-\epsilon)\tau}-e^{-4(\delta-\epsilon)\tau})\right]\lambda_t + \frac{3\alpha^4\theta\sigma^6}{4\kappa^3}(1-e^{-\kappa\tau})^4\\
&\quad + \frac{3\beta^4\delta\bar{\lambda}\epsilon^6}{4(\delta-\epsilon)^4}(1-e^{-(\delta-\epsilon)\tau})^4 + \frac{\beta^4\delta\bar{\lambda}\epsilon^4}{12(\delta-\epsilon)^2}(3-4e^{-(\delta-\epsilon)\tau}+e^{-4(\delta-\epsilon)\tau})\\
&\quad + \frac{\beta^4\delta\bar{\lambda}\epsilon^5}{3(\delta-\epsilon)^3}\left(3-7e^{-(\delta-\epsilon)\tau}+3e^{-2(\delta-\epsilon)\tau}+3e^{-3(\delta-\epsilon)\tau}-2e^{-4(\delta-\epsilon)\tau}\right).
\end{aligned}$$

Proof. With the conditional characteristic function given in Proposition 3.1, we get the cumulants of $VIX_T'^2$ as follows:

$$c_j = \frac{1}{i^j}\left(\frac{\partial^j A(\phi;\tau)}{\partial\phi^j}\bigg|_{\phi=0}V_t + \frac{\partial^j B(\phi;\tau)}{\partial\phi^j}\bigg|_{\phi=0}\lambda_t + \frac{\partial^j C(\phi;\tau)}{\partial\phi^j}\bigg|_{\phi=0}\right).$$

For the sake of brevity, we focus on the derivation of the first cumulant. Taking the derivative with respect to ϕ in the above system of ODEs yields:

$$\begin{aligned}
\frac{\partial}{\partial\tau}\left(\left.\frac{\partial A(\phi;\tau)}{\partial\phi}\right|_{\phi=0}\right) &= -\kappa\left.\frac{\partial A(\phi;\tau)}{\partial\phi}\right|_{\phi=0} + \sigma^2 A(\phi;\tau)\left.\frac{\partial A(\phi;\tau)}{\partial\phi}\right|_{\phi=0},\\
\frac{\partial}{\partial\tau}\left(\left.\frac{\partial B(\phi;\tau)}{\partial\phi}\right|_{\phi=0}\right) &= -\delta\left.\frac{\partial B(\phi;\tau)}{\partial\phi}\right|_{\phi=0} + \epsilon e^{\epsilon B(\phi;\tau)}\left.\frac{\partial B(\phi;\tau)}{\partial\phi}\right|_{\phi=0},\\
\frac{\partial}{\partial\tau}\left(\left.\frac{\partial C(\phi;\tau)}{\partial\phi}\right|_{\phi=0}\right) &= \kappa\theta\left.\frac{\partial A(\phi;\tau)}{\partial\phi}\right|_{\phi=0} + \delta\bar{\lambda}\left.\frac{\partial B(\phi;\tau)}{\partial\phi}\right|_{\phi=0},
\end{aligned}$$

subject to the boundary conditions $\left.\frac{\partial A(\phi;0)}{\partial\phi}\right|_{\phi=0} = i\alpha$, $\left.\frac{\partial B(\phi;0)}{\partial\phi}\right|_{\phi=0} = i\beta$ and $\left.\frac{\partial C(\phi;0)}{\partial\phi}\right|_{\phi=0} = i\gamma$. Note that $A(0;\tau)$, $B(0;\tau)$ and $C(0;\tau)$ are all equal to zero. This system can be simplified as:

$$\begin{aligned}
\frac{\partial}{\partial\tau}\left(\left.\frac{\partial A(\phi;\tau)}{\partial\phi}\right|_{\phi=0}\right) &= -\kappa\left.\frac{\partial A(\phi;\tau)}{\partial\phi}\right|_{\phi=0},\\
\frac{\partial}{\partial\tau}\left(\left.\frac{\partial B(\phi;\tau)}{\partial\phi}\right|_{\phi=0}\right) &= -(\delta-\epsilon)\left.\frac{\partial B(\phi;\tau)}{\partial\phi}\right|_{\phi=0},\\
\frac{\partial}{\partial\tau}\left(\left.\frac{\partial C(\phi;\tau)}{\partial\phi}\right|_{\phi=0}\right) &= \kappa\theta\left.\frac{\partial A(\phi;\tau)}{\partial\phi}\right|_{\phi=0} + \delta\bar{\lambda}\left.\frac{\partial B(\phi;\tau)}{\partial\phi}\right|_{\phi=0}.
\end{aligned}$$

The above system of ODEs can be solved analytically as follows:

$$\begin{aligned}
\left.\frac{\partial A(\phi;\tau)}{\partial\phi}\right|_{\phi=0} &= i\alpha e^{-\kappa\tau},\\
\left.\frac{\partial B(\phi;\tau)}{\partial\phi}\right|_{\phi=0} &= i\beta e^{-(\delta-\epsilon)\tau},\\
\left.\frac{\partial C(\phi;\tau)}{\partial\phi}\right|_{\phi=0} &= i\alpha\theta(1-e^{-\kappa\tau}) + \frac{i\beta\delta\bar{\lambda}}{\delta-\epsilon}(1-e^{-(\delta-\epsilon)\tau}) + i\gamma.
\end{aligned}$$

Therefore, the first cumulant is given by:

$$c_1 = \alpha e^{-\kappa\tau}V_t + \beta e^{-(\delta-\epsilon)\tau}\lambda_t + \alpha\theta(1-e^{-\kappa\tau}) + \frac{\beta\delta\bar{\lambda}}{\delta-\epsilon}(1-e^{-(\delta-\epsilon)\tau}) + \gamma.$$

For the second and fourth cumulants, they can be derived in an analogous manner. □

With the two lemmas above, we can state the following proposition without proof.

Proposition 4.3. *Under the SV-HJ model described in (1), given the truncation interval $[a,b]$, the time-t price of a VIX call option with strike price K and maturity T is approximated by:*

$$C_t \approx 100e^{-r\tau}\sum_{n=0}^{N-1}{}'\Re\left[\Phi\left(\frac{n\pi}{b-a};V_t,\lambda_t,\tau\right)e^{-i\frac{na\pi}{b-a}}\right]U_n, \tag{14}$$

where $\Phi(\cdot)$ is the conditional characteristic function given in Proposition 3.1 and the coefficients U_n are given in Lemma 4.1.

For a VIX futures contract with maturity time T, its price at time t, denoted by F_t, can be expressed as:

$$F_t = E_t^{\mathbb{Q}}\left(VIX_T\right) = 100E_t^{\mathbb{Q}}\left(\sqrt{VIX_T'^2}\right). \tag{15}$$

Working through an analogous derivation as above, we can obtain a pricing formula for the VIX futures based on the COS method, as presented in the following corollary.

Corollary 4.4. *Under the SV-HJ model described in (1), given the truncation interval [a,b], the time-t price of a VIX futures contract with maturity time T is approximated by:*

$$F_t \approx 100\sum_{n=0}^{N-1}{}'\Re\left[\Phi\left(\frac{n\pi}{b-a};V_t,\lambda_t,\tau\right)e^{-i\frac{na\pi}{b-a}}\right]\tilde{U}_n, \tag{16}$$

where $\Phi(\cdot)$ is the conditional characteristic function given in Proposition 3.1 and $\tilde{U}_n$ are the Fourier cosine series coefficients of payoff $\sqrt{y}$ given as follows:

$$\tilde{U}_n = \begin{cases} \frac{4}{3(b-a)}(b^{\frac{3}{2}}-a^{\frac{3}{2}}), & n=0,\\ \frac{2}{b-a}\Re\left\{e^{-i\omega_n a}\left[\frac{\sqrt{b}}{i\omega_n}e^{i\omega_n b} - \frac{\sqrt{a}}{i\omega_n}e^{i\omega_n a} + \frac{\sqrt{\pi}}{2(\sqrt{-i\omega_n})^3}\right.\right. & \\ \left.\left.\cdot\left(erfz(\sqrt{-i\omega_n b}) - erfz(\sqrt{-i\omega_n a})\right)\right]\right\}, & n\neq 0. \end{cases}$$

Table 1
Numerical pricing results of VIX options.

K = 15	Price		CPU time	
DTM	COS	MC	COS	MC
60	15.0343	14.9781	0.1549	5.1399
90	15.2068	15.2103	0.1508	4.7844
120	15.3094	15.3347	0.1305	5.3405
K = 20	Price		CPU time	
DTM	COS	MC	COS	MC
60	9.9294	9.9575	0.1241	5.4217
90	10.2444	10.2689	0.1127	4.3286
120	10.4325	10.4279	0.1098	4.1357

Table 1 shows the prices of VIX call options with different strike prices and maturities calculated from the COS method and MC simulations with the parameters $(\kappa,\ \theta,\ \sigma,\ V_0,\ \delta,\ \bar{\lambda},\ \epsilon,\ \lambda_0,\ \mu_s,\ \sigma_s)$ = (4.9363, 0.0672, 1.3837, 0.0046, 7.7198, 0.5992, 2.6118, 1.5026, −0.0892, 0.2082) corresponding to a set of calibrated parameters below. The number of sample paths reaches 100,000 in the MC simulations. DTM denotes the number of calendar days to maturity and CPU time is expressed in seconds. Computations are done on a laptop running macOS Sierra 10.12 with a 2.3 GHz Intel Core i5 processor and 8 GB of RAM.

Proof. The proof of this corollary is straightforward by setting the interest rate and strike price to zero in Lemma 4.1 and Proposition 4.3. □

Note that the price of a put option can be derived analogously. Alternatively, one can use the modified put–call parity to obtain the put option pricing formula, which is given by:

$$P_t = C_t - e^{-r(T-t)}\left(F_t - K\right) \tag{17}$$

The modified put–call parity differs from the traditional one in that the underlying price is replaced by the discounted forward volatility due to the lack of the typical cost-of-carry relation between spot VIX and VIX futures (see Lian & Zhu, 2013; Shu & Zhang, 2012; Wang & Daigler, 2011).

5. Numerical experiments

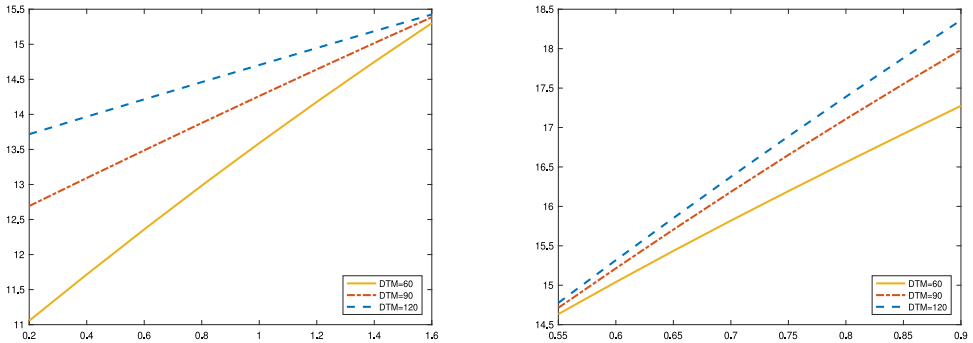
We first provide several numerical examples to demonstrate the reliability and efficiency of the COS method for pricing VIX call options by comparing with the pricing results of MC simulations. As for the MC simulations, we adopt the simple Euler–Maruyama discretization for the SV-HJ model, in which the Hawkes process N_t is simulated using the thinning algorithm proposed by Ogata (1981). With the simulated paths of V_t and λ_t, the value of VIX can be calculated according to (4). As a result, the prices of VIX options can be obtained based on this simulation procedure. Table 1 presents the numerical results of VIX options with varying strike prices and maturities calculated from the COS method and MC simulations. We also check the computational efficiency by comparing the CPU time required by these two methods. It can be seen that there are no significant differences between the numerical results obtained from the COS method and MC simulations. The good agreement of the numerical pricing results confirms the validity and reliability of the approximation method. Moreover, the COS method has the benefit in terms of computational efficiency, which requires much less CPU time than the MC simulations.

Further, in order to explore the impact of jump clustering on option prices, we then perform a sensitivity analysis with changing one of the parameters at a time while keeping the values of the remaining parameters unchanged. Fig. 2 shows the sensitivity of the option prices with respect to the jump clustering related parameters for varying maturities. We observe that the initial jump intensity λ_0 and the background intensity $\bar{\lambda}$ have similar results with impacting positively the option values. It is also clear that the option prices increase significantly with the increase of the intensity-based jump-size parameter, ϵ, as shown in Fig. 2c. On the contrary, the decay rate of jump clustering δ is negatively related to the option prices. It is interesting to note that options with longer maturity are more sensitive to the jump clustering related parameters. Overall, the above results of these numerical experiments illustrate that our approximation pricing formula for VIX options is reliable and efficient, and jump clustering has a significant impact on VIX option values.

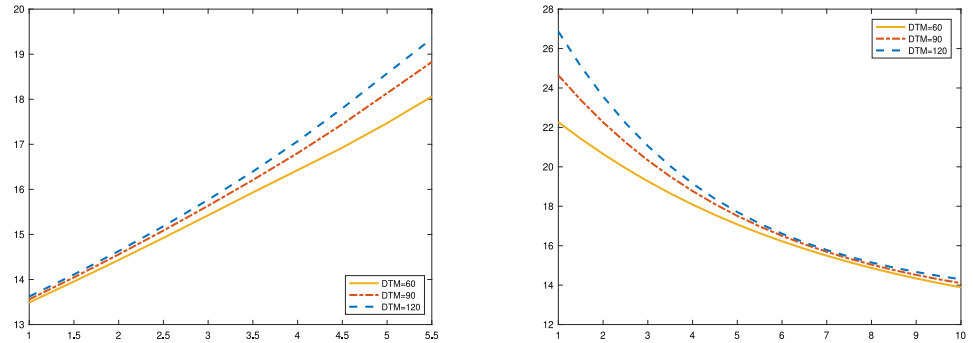
6. Empirical studies

6.1. Data description

The primary market data used for empirical studies are collected from OptionMetrics. Kokholm and Stisen (2015) selected two days, i.e., October 22, 2008 and May 16, 2012, to compare the pricing performance in distress and regular market environments. To obtain the average results, we calibrate the models to the entire cross-sectional options data for two months October 2008 and May 2012, which cover the above two days. By convention, the mid-quote, the average of bids and asks, is used as the market price of the VIX option. To avoid noise in the data, we apply several commonly adopted filtering rules to the primary data as is typically



(a) Option prices against λ_0 (b) Option prices against $\bar{\lambda}$



(c) Option prices against ϵ (d) Option prices against δ

Fig. 2. VIX option prices with $K = 15$ for different maturities against different parameters.

Table 2
Descriptive statistics of the VIX options data.

	Observations	Mean	Maximum	Minimum	Std	Skewness	Kurtosis
October 2008	1539	8.62	49.25	0.10	8.94	1.54	2.24
May 2012	2200	3.71	18.70	0.08	3.45	1.38	2.02

The table provides descriptive statistics of the VIX options data on October 2008 and May 2012. Std is the abbreviation of standard deviation. Kurtosis is the excess kurtosis.

done in the literature. Specifically, all options with fewer than 7 or more than 365 calendar days to maturity are excluded due to their illiquidity. Moreover, we remove option quotes with bid prices equal to zero. Finally, options whose open interests and volume are equal to zero are filtered out as well. Table 2 presents the descriptive statistics of the VIX options data. As expected, the option prices during a turbulent period are more volatile, with the standard deviation being 8.94.

6.2. Empirical results

To illustrate the ability of our jump specification to fit the market data, we consider the Bates-type jump intensity specification as a benchmark for comparison. Bates (2000), Eraker (2004), and Pan (2002) used a linear specification $\lambda_0+\lambda_1 V_t$, for some non-negative constants λ_0 and λ_1, generating time-varying jump intensity. The specification implies that jumps arrive more frequently in periods with high volatility and depend on the dynamic specification adopted for variance itself over time. We refer to this model as the SV-BJ model. Using the filtered data described as above, we estimate the parameters of the models. Following Christoffersen and Jacobs (2004) and He and Zhu (2016), the model parameters are backed out by minimizing the mean-square error (MSE) between the market price and the model-determined price. The parameter estimation results over the two sample periods are summarized in Table 3. It should be noted that the estimated parameters vary substantially depending on the sample period irrespective of the model applied. For example, for the SV-HJ model, the mean-reversion speed is around 13.4771 based on the October 2008 data, which is larger than the parameter estimated from the May 2012 data. This is because the market term structure usually changes dramatically during a period of financial distress. The historical closing values of the VIX has remained at high levels during the extreme periods of the financial crisis. During a calm period of May 2012, the average closing value of the VIX is 21, while the average closing value is 61 for the period of October 2008. In addition, as mentioned earlier, the option prices during a turbulent

Table 3
Estimated parameters for the two models.

October 2008										
SV-HJ	κ	θ	σ	V_0	δ	$\bar{\lambda}$	ϵ	λ_0	μ_s	σ_s
	13.4771	0.0859	2.7154	0.4584	11.5038	0.8055	3.9538	3.4965	−0.0814	0.2017
SV-BJ	κ	θ	σ	V_0	λ_0	λ_1	μ_s	σ_s		
	13.2900	0.1071	2.4166	0.5311	1.7229	2.4898	−0.0089	0.0564		
May 2012										
SV-HJ	κ	θ	σ	V_0	δ	$\bar{\lambda}$	ϵ	λ_0	μ_s	σ_s
	4.9363	0.0672	1.3837	0.0046	7.7198	0.5992	2.6118	1.5026	−0.0892	0.2082
SV-BJ	κ	θ	σ	V_0	λ_0	λ_1	μ_s	σ_s		
	4.2825	0.0811	1.7543	0.0148	1.5777	2.4055	−0.0265	0.1153		

The table shows the estimated daily averaged parameters which are backed out by minimizing the mean-square error between the market price and the model-determined price.

Table 4
In-sample pricing errors for the two models.

October 2008	SV-HJ			SV-BJ		
DTM	MAE	RMSE	MAPE	MAE	RMSE	MAPE
Short-term	0.5428	0.7072	13.75	0.6063	0.7777	15.99
Medium-term	0.3136	0.4186	7.85	0.3308	0.4410	9.22
Long-term	0.2639	0.3809	11.78	0.2820	0.4355	9.79
All	0.4342	0.5935	11.74	0.4774	0.6504	13.09
May 2012	SV-HJ			SV-BJ		
DTM	MAE	RMSE	MAPE	MAE	RMSE	MAPE
Short-term	0.1302	0.1938	20.28	0.1372	0.1894	21.77
Medium-term	0.1292	0.1739	7.19	0.1636	0.2341	14.74
Long-term	0.1312	0.1943	6.30	0.1812	0.2831	11.14
All	0.1300	0.1867	12.52	0.1561	0.2282	16.96

The table shows the in-sample averaged pricing errors measured by the mean absolute error (MAE), the root mean-squared error (RMSE) and the mean absolute percentage error (MAPE) for the SV-BJ and SV-HJ models. The MAPEs are reported in percentage. VIX options are divided into three categories according to DTM. Short-term, Medium-term and Long-term options are defined by DTM ≤ 60, $60 \sim 120$ and ≥ 120, respectively.

period are more volatile. It suggests that the parameter estimation results may depend on the current economic condition. Therefore, it is necessary to examine the ability of the models to fit the VIX option surface in different market conditions.

In the following, with the estimated parameters obtained, the models are compared in the aspect of fitting the market data. To ensure the robustness of our results, three commonly used error measures are employed to evaluate the goodness of fit, i.e., the mean absolute error (MAE), the root mean-squared error (RMSE) and the mean absolute percentage error (MAPE), defined as:

$$\text{MAE} = \frac{1}{n_t}\sum_{i=1}^{n_t}\left|C_i^{Market} - C_i^{Model}\right|, \tag{18}$$

$$\text{RMSE} = \sqrt{\frac{1}{n_t}\sum_{i=1}^{n_t}\left(C_i^{Market} - C_i^{Model}\right)^2}, \tag{19}$$

$$\text{MAPE} = \frac{1}{n_t}\sum_{i=1}^{n_t}\frac{\left|C_i^{Market} - C_i^{Model}\right|}{C_i^{Market}}, \tag{20}$$

where n_t is the total number of contracts on day t, C_i^{Market} denotes the market price of contract i, C_i^{Model} represents the corresponding model-determined price. It is widely accepted that a model is regarded as the better one if it has less cumulative pricing errors. To assess forecasting abilities of the models for VIX options, we proceed to an out-of-sample pricing exercise with the previous day's parameters. Tables 4 and 5 report the in-sample and out-of-sample pricing errors of the models in terms of different measures, respectively.

It is clear that the pricing errors measured by the MAE and RMSE are larger during an extreme period of the financial crisis than those during a regular period, which holds for both SV-HJ and SV-BJ models. For example, for the SV-HJ model, the calibration error measured by the MAE increases from 0.1300 to 0.4342, indicating that the pricing performance of models deteriorates under extreme market conditions. It is interesting to note that the MAPEs from the 2008 data are basically smaller than those from the 2012 data. This is not difficult to understand because there are more expensive options (i.e., high market prices) during a turbulent period, as shown in Table 2, which results in lower MAPE values. From the perspective of in-sample pricing errors, the SV-HJ model achieves a better overall fit to the market data from May 2012 than the SV-BJ model. To be specific, the daily averaged RMSE for the SV-HJ model is 0.1867, compared with 0.2282 for the SV-BJ model. In other words, the SV-HJ model on average

Table 5
Out-of-sample pricing errors for the two models.

October 2008	SV-HJ			SV-BJ		
DTM	MAE	RMSE	MAPE	MAE	RMSE	MAPE
Short-term	0.9961	1.5349	17.91	0.9877	1.5061	19.16
Medium-term	0.5539	0.7846	12.68	0.5976	0.8139	15.81
Long-term	0.5353	0.8009	16.11	0.5380	0.7897	16.98
All	0.7982	1.2617	16.12	0.8065	1.2465	17.86
May 2012	SV-HJ			SV-BJ		
DTM	MAE	RMSE	MAPE	MAE	RMSE	MAPE
Short-term	0.3151	0.5637	22.72	0.3311	0.5601	28.21
Medium-term	0.2617	0.4003	11.31	0.2851	0.4189	18.47
Long-term	0.3176	0.4194	10.39	0.3078	0.4538	13.00
All	0.2956	0.4791	15.90	0.3091	0.4897	21.42

The table shows the out-of-sample averaged pricing errors measured by the mean absolute error (MAE), the root mean-squared error (RMSE) and the mean absolute percentage error (MAPE) for the SV-BJ and SV-HJ models. The MAPEs are reported in percentage. VIX options are divided into three categories according to DTM. Short-term, Medium-term and Long-term options are defined by DTM ≤ 60, $60 \sim 120$ and ≥ 120, respectively.

reduces the calibration errors over the SV-BJ model by 18.19%. Similar results are obtained for other error measures. Even under extreme market conditions, the SV-HJ model is still superior to the SV-BJ model, as it has lower pricing errors regardless of the error measures used. Turning the attention to out-of-sample case, it is not surprising that the out-of-sample pricing errors are always larger than in-sample pricing errors since model prices are calculated with the frozen parameters. We can see from Table 5 that, on the whole, the SV-HJ model has better predictive abilities than the SV-BJ model. For example, the out-of-sample errors measured by the MAE and MAPE for the SV-HJ model are 0.2956 and 15.90%, respectively, compared to their SV-BJ counterparts of 0.3091 and 21.42% in the May 2012 case. These consistent results illustrate the advantage of Hawkes jumps over Bates jumps in the aspect of fitting the market data. Further, taking a closer look at the in- and out-of-sample pricing performance for different categories according to maturity, we find that the SV-HJ model has lower pricing errors for most maturity categories, once again confirming that our jump specification provides a better pricing fit to the entire VIX option surface compared to the Bates-type jump. Putting the above empirical results together, we can conclude that our jump specification has a significant positive impact on the pricing performance.

7. Conclusion

In this paper, we employ a Hawkes jump-diffusion model to price VIX options in the consistent framework. Different from traditional jump-diffusion models, our proposed model is capable of capturing the clustering pattern of jumps consistent with what is observed in the financial markets. Under the Hawkes jump-diffusion model, which is analytically tractable due to its affine structure, we solve the pricing problem of VIX options based on the COS method. Finally, numerical experiments illustrate the reliability and efficiency of the pricing method and suggest that jump clustering has a significant impact on VIX option prices. To illustrate the ability of our jump specification to fit the market data, we offer a comparative study in which the Bates-type jump intensity specification is taken as a benchmark. Empirical results demonstrate the superiority of our jump specification in matching the VIX option surface, indicating the importance of accounting for jump clustering when pricing VIX options.

CRediT authorship contribution statement

Bo Jing: Methodology, Writing - original draft, Writing - review & editing, Investigation, Software. **Shenghong Li:** Supervision, Resources, Funding acquisition. **Yong Ma:** Conceptualization, Funding acquisition, Writing - review & editing, Software.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (grant numbers: 71971077 and 11571310).

References

Aït-Sahalia, Y., Cacho-Diaz, J., & Laeven, R. J. (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economic*, *117*(3), 585–606.

Bacry, E., & Muzy, J. F. (2014). Hawkes model for price and trades high-frequency dynamics. *Quantitative Finance*, *14*(7), 1147–1166.

Bardgett, C., Gourier, E., & Leippold, M. (2019). Inferring volatility dynamics and risk premia from the S&P 500 and VIX markets. *Journal of Financial Economic*, *131*(3), 593–618.

Bates, D. S. (2000). Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics*, *94*(1–2), 181–238.

Carr, P., & Madan, D. (1999). Option valuation using the fast fourier transform. *Journal of Computational Finance*, *2*(4), 61–73.

Chau, K.-W., Yam, S. C. P., & Yang, H. (2015). Fourier-cosine method for ruin probabilities. *Journal of Computational and Applied Mathematics*, *281*, 94–106.

Chavez-Demoulin, V., & McGill, J. A. (2012). High-frequency financial data modeling using Hawkes processes. *Journal of Banking & Finance*, *36*(12), 3415–3426.
Christoffersen, P., & Jacobs, K. (2004). Which GARCH model for option valuation? *Management Science*, *50*(9), 1204–1221.
Du, D., & Luo, D. (2019). The pricing of jump propagation: Evidence from spot and options markets. *Management Science*, *65*(5), 2360–2387.
Duan, J.-C., & Yeh, C.-Y. (2010). Jump and volatility risk premiums implied by VIX. *Journal of Economic Dynamics and Control*, *34*(11), 2232–2244.
Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, *68*(6), 1343–1376.
Eraker, B. (2004). Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *The Journal of Finance*, *59*(3), 1367–1403.
Errais, E., Giesecke, K., & Goldberg, L. R. (2010). Affine point processes and portfolio credit risk. *SIAM Journal on Financial Mathematics*, *1*(1), 642–665.
Fang, F., & Oosterlee, C. W. (2008). A novel pricing method for European options based on Fourier-cosine series expansions. *SIAM Journal on Scientific Computing*, *31*(2), 826–848.
Fulop, A., Li, J., & Yu, J. (2015). Self-exciting jumps, learning, and asset pricing implications. *Review of Financial Studies*, *28*(3), 876–912.
Goard, J., & Mazur, M. (2013). Stochastic volatility models and the pricing of VIX options. *Mathematical Finance*, *23*(3), 439–458.
Grünbichler, A., & Longstaff, F. A. (1996). Valuing futures and options on volatility. *Journal of Banking & Finance*, *20*(6), 985–1001.
Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, *58*(1), 83–90.
He, X., & Zhu, S. (2016). An analytical approximation formula for European option pricing under a new stochastic volatility model with regime-switching. *Journal of Economic Dynamics and Control*, *71*, 77–85.
Herrera, R., & Schipp, B. (2014). Statistics of extreme events in risk management: The impact of the subprime and global financial crisis on the German stock market. *The North American Journal of Economics and Finance*, *29*, 218–238.
Kokholm, T. (2016). Pricing and hedging of derivatives in contagious markets. *Journal of Banking & Finance*, *66*, 19–34.
Kokholm, T., & Stisen, M. (2015). Joint pricing of VIX and SPX options with stochastic volatility and jump models. *The Journal of Risk Finance*, *16*(1), 27–48.
Li, Z., Zhang, W., Zhang, Y., & Yi, Z. (2019). An analytical approximation approach for pricing European options in a two-price economy. *The North American Journal of Economics and Finance*, *50*, Article 100986.
Lian, G., & Zhu, S. (2013). Pricing VIX options with stochastic volatility and random jumps. *Decisions in Economics and Finance*, *36*(1), 71–88.
Lin, Y.-N. (2007). Pricing VIX futures: Evidence from integrated physical and risk-neutral probability measures. *Journal of Futures Markets*, *27*(12), 1175–1217.
Liu, W., & Zhu, S.-P. (2019). Pricing variance swaps under the Hawkes jump-diffusion process. *Journal of Futures Markets*, *39*(6), 635–655.
Luo, X., Zhang, J. E., & Zhang, W. (2019). Instantaneous squared VIX and VIX derivatives. *Journal of Futures Markets*, *39*(10), 1193–1213.
Ma, Y., Shrestha, K., & Xu, W. (2017). Pricing vulnerable options with jump clustering. *Journal of Futures Markets*, *37*(12), 1155–1178.
Mencía, J., & Sentana, E. (2013). Valuation of VIX derivatives. *Journal of Financial Economic*, *108*(2), 367–391.
Ogata, Y. (1981). On Lewis' simulation method for point processes. *IEEE Transactions on Information Theory*, *27*(1), 23–31.
Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*, *83*(401), 9–27.
Pacati, C., Pompa, G., & Renò, R. (2018). Smiling twice: The Heston++ model. *Journal of Banking & Finance*, *96*, 185–206.
Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics*, *63*(1), 3–50.
Park, Y. (2016). The effects of asymmetric volatility and jumps on the pricing of VIX derivatives. *Journal of Econometrics*, *192*(1), 313–328.
Psychoyios, D., Dotsis, G., & Markellos, R. N. (2010). A jump diffusion model for VIX volatility options and futures. *Review of Quantitative Finance and Accounting*, *35*(3), 245–269.
Shu, J., & Zhang, J. E. (2012). Causality in the VIX futures market. *Journal of Futures Markets*, *32*(1), 24–46.
Tour, G., Thakoor, N., Khaliq, A., & Tangman, D. (2018). COS method for option pricing under a regime-switching model with time-changed Lévy processes. *Quantitative Finance*, *18*(4), 673–692.
Wang, Z., & Daigler, R. T. (2011). The performance of VIX option pricing models: empirical evidence beyond simulation. *Journal of Futures Markets*, *31*(3), 251–281.
Zhang, B., & Oosterlee, C. (2014). Pricing of early-exercise Asian options under Lévy processes based on Fourier cosine expansions. *Applied Numerical Mathematics*, *78*, 14–30.
Zhang, X., Xiong, J., & Shen, Y. (2017). Bond and option pricing for interest rate model with clustering effects. *Quantitative Finance*, *18*(6), 969–981.
Zhang, J., & Zhu, Y. (2006). VIX futures. *Journal of Futures Markets*, *26*(6), 521–531.
Zhu, S.-P., & Lian, G.-H. (2012). An analytical formula for VIX futures and its applications. *Journal of Futures Markets*, *32*(2), 166–190.
Zhu, Y., & Zhang, J. E. (2007). Variance term structure and VIX futures pricing. *International Journal of Theoretical and Applied Finance*, *10*(01), 111–127.