Appl. Math. J. Chinese Univ.
2020, 35(1): 33-56

Pricing VIX options with stochastic skew and asymmetric jumps

JING Bo LI Sheng-hong TAN Xiao-yu

Abstract. This paper performs several empirical exercises to provide evidence that the stochastic skew behavior and asymmetric jumps exist in VIX markets. In order to adequately capture all of the features, we develop a general valuation model and obtain quasi-analytical solutions for pricing VIX options. In addition, we make comparative studies of alternative models to illustrate the effects after taking into account these features on the valuation of VIX options and investigate the relative value of an additional volatility factor and jump components. The empirical results indicate that the multi-factor volatility structure is vital to VIX option pricing due to providing more flexibility in the modeling of VIX dynamics, and the need for asymmetric jumps cannot be eliminated by an additional volatility factor.

§1 Introduction

Since the 2008 financial crisis, the trading volume of volatility derivatives has shown a trend of rapid growth. Among these diversified volatility derivatives, VIX options, as a kind of trading and hedging tools for a wide range of investors, have gained more and more popularity. This is partly because there is usually a negative relationship between stock returns and market volatilities, and therefore taking long positions in VIX call options can help investors to limit losses in a bear market environment. The research about VIX options introduced by the Chicago Board Options Exchange (CBOE) in 2006 has been attracting great attention in finance. In this sense, this paper focuses on the effects of stochastic skew and asymmetric jumps on the valuation of VIX options.

Numerous studies have been conducted on VIX options pricing in recent years (e.g., [13, 14, 16, 26]). It is well known that volatility has the significant mean-reverting effect. So far, the two most prominent mean-reverting models are the square root model (SR) and logarithmic model (LR). The latter first considered by [12] assumes that the logarithm of VIX follows an

Received: 2018-10-22. Revised: 2019-10-23.
MR Subject Classification: 62P05, 91B25, 91G20.
Keywords: VIX options, stochastic skew, multi-factor volatility, jumps.
Digital Object Identifier(DOI): https://doi.org/10.1007/s11766-020-3650-0.
Supported by the National Natural Science Foundation of China (11571310, 71371168).

Ornstein-Uhlenbeck (OU) process. [6] extended the LR model to incorporate upward jumps and concluded that the LR model serves better than the SR model in fitting VIX historical data and VIX option quotes. [13] led to the same conclusion as well. On the basis of the wide acceptance of the LR model in VIX literature relative to the SR model, we would like to follow the LR framework to model the dynamics of the VIX index.

It is widely recognized that the implied volatility surface extracted by the Black−Scholes formula displays a smiling pattern in options markets. To account for this feature, a variety of models have been proposed, including the stochastic volatility (SV) models. In contrast with the downward implied volatility skew observed in the S&P 500 index (SPX) options markets, the implied volatility smiles of VIX options display upward sloping shape, found by [2]. [14] developed an analytical pricing formula for VIX options under the 3/2 model and pointed out that the model is capable of generating the upward implied volatility skew. Nevertheless, the strong time-variation in the slope of the implied volatility smile poses a new modeling challenge for option pricing theory. There are several empirical studies providing evidence that the slopes of the implied volatility smiles of foreign exchange options and equity index options are time varying (e.g., [19, 21]). The feature that the skew evolves stochastically through time observed in the VIX options markets is first illustrated by [15]. As far as we know, this stylized fact is not extensively documented elsewhere. To this end, we perform several empirical exercises to provide additional evidence via a quadratic polynomial fit on the implied volatility smiles.

As pointed out by [19], there are two approaches to capture the stochastic skew feature, the first one is to randomize the mean jump size and the second one is to randomize the correlation between the asset return and its instantaneous variance process. In this paper, we take the latter approach to allow for stochastic correlation by considering multi-factor volatility models. Starting with [25], lots of option pricing literature found that multi-factor volatility models are preferred to single-factor ones (e.g., [5, 20]). Based on a principal component analysis of the implied variance, [21] documented that the first two components together explain more than 95% of the variation in the data. Then they applied a two-factor Heston model, which displays stochastic correlation between the noises driving the stock return and the stochastic volatility process, to SPX options pricing. Within the single-factor framework, on the contrary, the correlation is constant over time which limits the ability of capturing stochastic skew. Also, [17] indicated that the introduction of an additional stochastic volatility factor improves the pricing performance of target volatility options.

Apart from stochastic skew, jumps are widely regarded as a salient feature of volatility. [18] found evidence of significant upward jumps in implied volatilities. By comparing jump diffusion and continuous-time diffusion processes with respect to the ability of capturing the dynamics of volatility indices, [9] concluded that the former is the best fit to the data. It is well known that good and bad surprises will cause upward and downward jumps in stock prices. Since the Merton jump component, where jump size follows the normal distribution, has a limitation on distinguishing the type of jumps, the double exponential jump-diffusion model is proposed by [24] which can differentiate upward and downward jumps by employing different exponential distributions. Hence, in order to study the impact of asymmetric jumps on the pricing of VIX options, we adopt the double exponential specification. Furthermore, we consider the negative exponential specification that only captures upward jumps in the VIX dynamics to analyze whether including downward jumps offers benefits for VIX option valuation in the context of multi-factor volatility models.

Motivated by the critical empirical features mentioned above, we develop a general valuation model which captures these features by employing the multi-factor volatility structure and the double exponential jump diffusion process. To the best of our knowledge, most studies on VIX option pricing do not take stochastic skew into account. For example, [14] employed the 3/2 model to price VIX options, but this model is incapable of capturing this feature. Although [15] first considered stochastic skew in the valuation of volatility options, the author did not consider jump components nor test the pricing improvements for the additional volatility factor. [4] also investigated the pricing performance associated with VIX options across different models. However, we differ from [4] by adopting the standalone approach with modeling the VIX dynamics directly.

This paper mainly contributes to the existing literature in several ways. First, the proposed model is a general model, which adequately captures critical modeling features. It can be regarded as an extension of the models proposed by [3, 6, 15]. Second, we find quasi-analytical expressions for the characteristic function and thereby solving the pricing problem of VIX derivatives. Finally, within the standalone framework, we make comparative studies of different option valuation models to analyze the in-sample and out-of-sample pricing performance. To further investigate the marginal contributions for the additional components in explaining the option prices, we also test the pricing improvements across different models.

The remainder of this paper proceeds as follows. Section 2 performs empirical studies to document the main features of VIX dynamics. Section 3 displays the model framework and solves the pricing problem associated with VIX derivatives. The calibration procedure is conducted in Section 4. Section 5 discusses the empirical results of the pricing performance across different model specifications. This paper is concluded in Section 6.

§2 Empirical Analyses

2.1 Data description

In 1993, the Chicago Board Options Exchange (CBOE) introduced a volatility index, the VIX index, originally designed to measure the implied 30-day volatility of the S&P 100 index. After September 22, 2003, with the new methodology adopted, the VIX index reflects the expected market volatility implied by SPX options over the next month, and the index using the old methodology was renamed to VXO. Following the change from the old VIX (VXO) to the new VIX, the CBOE launched VIX futures and options as tradable assets on March 26, 2004 and February 24, 2006, respectively. Our primary dataset comprises the daily time-series of VIX and VIX option quotes obtained from the website of CBOE and OptionMetrics. The

Table 1: Summary statistics for VIX option data.

Panel A: Number of option contracts

	TTM≤60	60<TTM≤120	TTM>120	All
$m \leq -0.1$	638	261	210	1,109
$-0.1 < m \leq 0.1$	1,591	466	395	2,452
$0.1 < m \leq 0.3$	1,805	579	438	2,822
$0.3 < m \leq 0.5$	1,528	687	508	2,723
$m > 0.5$	3,071	2,756	2,256	8,083
All	8,633	4,749	3,807	17,189

Panel B: Average option prices

	TTM≤60	60<TTM≤120	TTM>120	All
$m \leq -0.1$	3.11	4.71	5.80	4.00
$-0.1 < m \leq 0.1$	1.81	3.66	4.85	2.65
$0.1 < m \leq 0.3$	0.94	2.41	3.45	1.63
$0.3 < m \leq 0.5$	0.53	1.59	2.34	1.13
$m > 0.5$	0.25	0.65	0.96	0.58
All	0.94	1.52	2.10	1.36

Panel C: Average implied volatility

	TTM≤60	60<TTM≤120	TTM>120	All
$m \leq -0.1$	0.80	0.54	0.45	0.65
$-0.1 < m \leq 0.1$	0.74	0.57	0.51	0.67
$0.1 < m \leq 0.3$	0.96	0.71	0.60	0.85
$0.3 < m \leq 0.5$	1.22	0.85	0.69	1.03
$m > 0.5$	1.52	1.06	0.83	1.17
All	1.17	0.92	0.73	1.00

This table reports the summary statistics for VIX call options data, which are obtained from OptionMetrics, during the period from January 2, 2017 to December 31, 2017. The VIX option dataset is divided into five moneyness categories and three time-to-maturity (TTM) categories, where the moneyness is defined as $m = \log(\frac{K}{S})$, with K the strike level of options and S the VIX closing price. To avoid noise in the dataset, options with time-to-maturity fewer than seven days and zero bids are deleted. In addition, options whose open interests and volume are equal to zero are filtered out as well. The number of option contracts, the average option prices and the average implied volatilities are listed in the Panel A, B and C, respectively.

VIX index closing values from January 2, 2004 through July 17, 2018 are used to investigate empirical properties of the VIX dynamics. To document the features of market volatility implied by VIX options, we use the end-of-day option quotes during the period from January 2, 2017 to December 31, 2017, spanning one year.

Following the standard convention in the literature, the mid-prices defined as the average of bids and asks are the proxy for the option prices. To avoid noise in the dataset, we adopt several filtering rules. Specifically, all options with time-to-maturity fewer than seven days and zero bids are deleted. In addition, options whose open interests and volume are equal to zero are filtered out as well. Finally, following [22], we use Wednesday option data as in-sample data and Thursdays option data as out-of-sample data to avoid the day-of-the-week effect. After applying these rules for selecting option data, a total of 17,189 observations remain in our sample. Table 1 summarizes VIX option filtered dataset divided into five categories according to moneyness and three categories based on time-to-maturity (TTM) in daily units. In this paper, the moneyness is defined as $m = \log(\frac{K}{S})$, where K is the strike level of options and S denotes the VIX closing price. This definition implies that a call option is out-of-the-money (OTM) when m value is positive and a larger value of m means a deeper OTM option. Note that the number of OTM option contracts is greater than the number of ITM options, probably because investors are afraid of a stock market crash and prefer to buy OTM options when market prospects are uncertain. Looking across each column of Panel C, the implied volatility smiles of VIX options display upward skew pattern consistent with empirical evidence found by [2, 10], among others.

2.2 Empirical features

Using the daily time-series of VIX and VIX option quotes, we document several important features that a VIX option pricing model should capture. Figure 1 depicts the daily VIX logarithmic returns during the sample period. The plot suggests a mean-reversion (MR) nature for VIX log-returns. We also observe that there are several large movements.

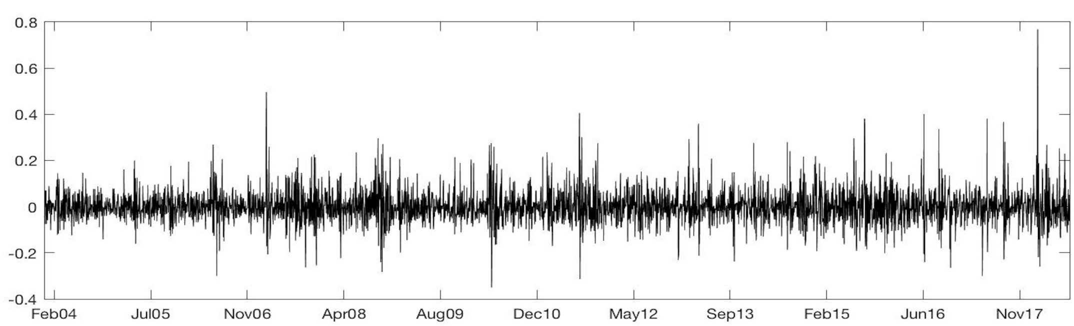


Figure 1: Daily logarithmic returns on the VIX index from January 2, 2004 until July 17, 2018.

To further dissect this feature, the descriptive statistics for changes in logarithmic VIX are provided in Table 2. The skewness and the kurtosis in returns are respectively 0.9953 and 7.117, suggesting that the return distribution is skewed to the right and has the leptokurtic feature. Given that the number of samples is 3,660, the standard deviation is 0.0722, and the mean is around 0, there are 19 observations exceeding the range of four-standard deviations: 15 above this range and 4 below this range, which account for approximately 0.52% of the total. Whereas, the probability of variations happening is under 0.0063% in a normal distribution. These findings confirm the existence of sudden jumps and heavier tails in the distribution of VIX log-returns. Accordingly, jumps as a salient feature of VIX dynamics should be incorporated in an option pricing model. Besides, we find that although jumps usually occur to be upward, there are still some downward jumps. In this case, it is necessary to consider asymmetric jumps and investigate the effect of including downward jumps on option pricing performance.

Table 2: Descriptive statistics for VIX log-returns.

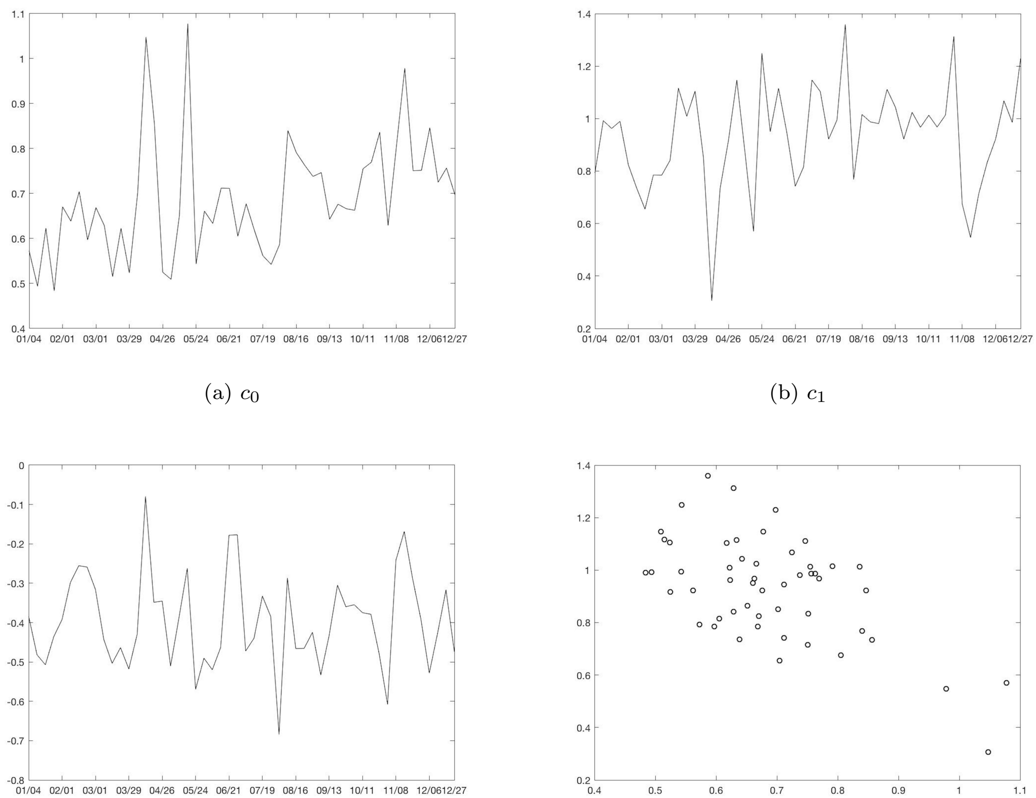
	Observations	Mean	Median	Max	Min	Std	Skew	Kurt
VIX log-returns	3,660	-0.0001	-0.0054	0.7682	-0.3506	0.0722	0.9953	7.1170

This table provides descriptive statistics for VIX log-returns from January 2, 2004 until July 17, 2018. Std is the abbreviation of standard deviation. Kurt is the excess kurtosis.

As said previously, the stochastic skew feature that the slope of the implied volatility smile varies greatly over time has been documented by some empirical studies in equity options markets and foreign exchange options markets. However, to the best of our knowledge, there are little literature documenting this feature with respect to the VIX index. In order to illustrate the existence of stochastic skew in VIX options markets as well, we fit the implied volatility smiles through a quadratic polynomial in moneyness:

$$\sigma = c_0 + c_1 m + c_2 m^2 + e, \tag{1}$$

where σ is the implied volatility, m is the previously defined moneyness and e is the error term following normal distribution with mean zero. In this way, the intercept c_0 can be interpreted as the level of the at-the-money (m=0) implied volatility, and the other two coefficients c_1 and c_2 measure the slope and the curvature of the smile, respectively. This form is an extension of the framework proposed in [21] by incorporating a quadratic term in the implied volatility function. Under this specification, we consider not only the slope of the smile, but also the curvature of the smile characterized by the coefficient for squared moneyness c_2, which integrates all the information contained in the implied volatility smile. Adding a quadratic term can better describe the shape of the implied volatility smile. Therefore, we employ a quadratic polynomial in moneyness to fit the smiles more precisely, which, in turn, ensures the accurate measurement of the slope of the smile. In this way, we can well demonstrate the existence of stochastic skew in the VIX options markets.



(a) c_0

(b) c_1

(c) c_2

(d) c_1 v.s. c_0

Figure 2: Time series of c_0, c_1 and c_2 and scatter plot of c_1 v.s. c_0.

The regression in Equation (1) is estimated on each Wednesday during the sample period and then we can obtain the time series of resulting coefficients which are depicted in Figure 2. It is obvious from the top two panels that both the volatility level and the slope show substantial time variation during our sample period. Moreover, the values of c_1 are larger than 0, confirming the stylized fact that the implied volatility smiles of VIX options display upward skew pattern. Thus, an option pricing model should be capable of accounting for stochastic volatility, upward volatility skew and stochastic skew observed in the VIX options markets. In this paper, we adopt the multi-factor volatility models as the means of capturing stochastic skew. Another reason for considering the multi-factor volatility structure is that models without a second volatility factor have a limitation on explaining independent movements in the level and slope of option-implied volatility smile. From the visual point of view, it seems that the volatility level is independent of the skew in the top two panels of Figure 2. To further verify this, a scatter plot of c_1 against c_0 is displayed in the bottom right panel of Figure 2. It reveals that there are steep and flat slopes of the smile regardless of the levels of implied volatility.

Based on aforementioned empirical analyses, a sound VIX option pricing model should be able to adequately capture these important features.

§3 Multi-factor Jump Diffusions

3.1 The model framework

In order to adequately capture all the features previously mentioned, we develop a general valuation model which is characterized by the considerations on two distinctive features, stochastic skew and asymmetric jumps. Then, we specify the four nested models so as to analyze the marginal contributions for the additional components in explaining the option prices later. We employ the multi-factor volatility structure and the double exponential jump diffusion process as the means of capturing stochastic skew and asymmetric jumps. Given a complete probability space $(\Omega,\ \mathcal{F},\ \mathbb{Q})$, equipped with an information filtration $\{\mathcal{F}_t\}_{t\geq 0}$, where $\mathbb{Q}$ is the risk-neutral probability measure. On the probability space $(\Omega,\ \mathcal{F},\ \mathbb{Q})$, the dynamics for logarithmic underlying asset price, denoted by $X_t = \log VIX_t$, and its instantaneous variance processes jointly evolve as follows:

$$\begin{aligned} dX_t &= k\left(\theta - X_t\right)dt + \sqrt{V_{1t}}dW_t^1 + \sqrt{V_{2t}}dW_t^2 + Y dN_t. \\ dV_{1t} &= k_1\left(\theta_1 - V_{1t}\right)dt + \sigma_1\sqrt{V_{1t}}\left(\rho_1 dW_t^1 + \sqrt{1-\rho_1^2}dW_t^3\right) \\ dV_{2t} &= k_2\left(\theta_2 - V_{2t}\right)dt + \sigma_2\sqrt{V_{2t}}\left(\rho_2 dW_t^2 + \sqrt{1-\rho_2^2}dW_t^4\right) \end{aligned} \tag{2}$$

where W_t^1, W_t^2, W_t^3 and W_t^4 refer to independent standard Brownian motions; k and θ are respectively the mean-reversion speed and the long-run mean level of X_t; k_i, θ_i and σ_i respectively represent the mean-reversion speed, the long-run mean and the volatility of the ith instantaneous variance factor V_{it}, $i = 1, 2$; N_t denotes Poisson process with constant intensity λ; and we assume the jump size Y to be drawn from the asymmetric double exponential distribution with the density:

$$f(y) = p \cdot \eta_1 e^{-\eta_1 y}\mathbb{I}_{\{y\geq 0\}} + q \cdot \eta_2 e^{\eta_2 y}\mathbb{I}_{\{y<0\}}, \tag{3}$$

where $\eta_1 > 1$, $\eta_2 > 0$, $p, q \geq 0$ and $p + q = 1$. The values of p and q respectively represent the probability of upward jumps and downward jumps. The values of $1/\eta_1$ and $1/\eta_2$ stand for the average size of upward jumps and downward jumps. We refer to this general model as the MSV-AJ model.

There are four option pricing models nested within the above-mentioned model framework. The first two models, denoted by SSV and SSV-UJ, are both the single-factor stochastic volatility models, while the difference between them is that the latter, proposed by [3], allows for upward jumps. If the jump component is removed from the general specification, the model is reduced to a multi-factor stochastic volatility model introduced by [15], referred to as the MSV

Table 3: Summary of model specifications.

Model	Restrictions on the parameters	Description
SSV	k_2=θ_2=σ_2=ρ_2=V_{20}=0, λ=0	Single-factor stochastic volatility without jumps
SSV-UJ	k_2=θ_2=σ_2=ρ_2=V_{20}=0, q=0	Single-factor stochastic volatility with upward jumps
MSV	λ=0	Multi-factor stochastic volatility without jumps
MSV-UJ	q=0	Multi-factor stochastic volatility with upward jumps
MSV-AJ	Not applicable	Multi-factor stochastic volatility with asymmetric jumps

This table shows different model specifications discussed in this paper. SSV denotes the single-factor stochastic volatility model with no jumps. SSV-UJ adds upward jumps in VIX in the SSV model. MSV introduces an additional volatility factor in the SSV model. MSV-UJ adds upward jumps in VIX in the MSV model. MSV-AJ adds asymmetric jumps in VIX in the MSV model.

model, which can capture stochastic skew. Finally, setting $q = 0$ (i.e. $p = 1$) yields negative exponential distribution used by [6] to capture only upward jumps. The fourth nested model labeled MSV-UJ is adopted to investigate whether incorporating downward jumps improves the pricing performance associated with VIX options in the context of the multi-factor volatility structure. All model specifications to be discussed are summarized in Table 3.

Compared with the single-factor models, the multi-factor volatility structure is able to generate stochastic correlation between the noises driving the stock return and the stochastic volatility process, whereas the single-factor models cannot. Based on the discussions of [19, 21], this property enables the multi-factor models to capture stochastic skew consistent with the empirical evidence. On the other hand, jumps in the underlying asset are modeled by the double exponential specification which allows for asymmetric jumps and generates a highly skewed and leptokurtic distribution matching the market data. In a word, specifying the multi-factor volatility structure and introducing the double exponential specification both provide more flexibility than other models in the modeling of VIX dynamics. Therefore, theoretically, the multi-factor model with asymmetric jumps may have better explanatory power to the market data than other models.

3.2 Pricing VIX options

Now we turn to deal with the pricing problem of VIX options. Due to the complexity of the general specification (2), an analytical pricing formula is not directly derived and therefore we take the characteristic function approach. The conditional characteristic function of X_t under the risk-neutral measure $\mathbb{Q}$ is defined as:

$$g\left(X_t, \tau; s\right) := E_t^{\mathbb{Q}}\left[exp(isX_T)\right], \tag{4}$$

where $\tau = T - t$, $E_t^{\mathbb{Q}}[\cdot]$ denotes the expectation conditional on the information available up to time t under the $\mathbb{Q}$ measure and $i^2{=}{-1}$. We then have the following Lemma, which is an extension of [15].

Lemma 1. *For the MSV-AJ model (2), the conditional characteristic function defined in (4) is given by:*

$$g\left(X_t, \tau; s\right) = exp[A(\tau; s) + \sum_{i=1}^{2} B_i(\tau; s) V_{it} + C(\tau; s) X_t], \tag{5}$$

where $A(\tau; s)$, $B_i(\tau; s)$, i=1, 2 and $C(\tau; s)$ can be obtained by solving the following ODE system:

$$\begin{aligned} \frac{\partial A}{\partial \tau} &= k\theta C + \sum_{i=1}^{2} k_i \theta_i B_i + \lambda(\frac{p\eta_1}{\eta_1 - C} + \frac{q\eta_2}{\eta_2 + C} - 1) \\ \frac{\partial B_i}{\partial \tau} &= \frac{1}{2}C^2 + (C\rho_i\sigma_i - k_i)B_i + \frac{1}{2}B_i^2\sigma_i^2 \\ \frac{\partial C}{\partial \tau} &= -kC \end{aligned} \tag{6}$$

with initial conditions $A\left(0; s\right) = 0$, $B_i\left(0; s\right) = 0$, $i = 1, 2$ and $C\left(0; s\right) = is$.

Proof. See Appendix A. □

We remark here that the coefficient functions $B_i(\tau;$ s) in the above ODE system known as the Riccati equations can be solved by using numerical method, the Runge−Kutta algorithm, in our experiment to improve calculation speed as well as retain accuracy. Once the characteristic function is found, we can directly derive the pricing formulas for VIX futures and VIX options.

Lemma 2. *For the MSV-AJ model (2), the time-t price of VIX futures with maturity T is given by*

$$F_t^T = g\left(X_t, \tau; -i\right). \tag{7}$$

Proof. Using the characteristic function and risk-neutral valuation theory, it is trivial to obtain the pricing formula for VIX future:

$$F_t^T = E_t^{\mathbb{Q}}\left(VIX_T\right) = E_t^{\mathbb{Q}}(e^{X_T}) = g\left(X_t, \tau; -i\right)$$ □

Theorem 1. *For the MSV-AJ model (2), the time-t prices of European-style call options and put options written on VIX with strike price K and maturity T, denoted $C_t^T(VIX_t, K)$ and $P_t^T(VIX_t, K)$ respectively, are given by*

$$C_t^T\left(VIX_t, K\right) = e^{-r(T-t)}\left[F_t^T \cdot \Pi_1 - K \cdot \Pi_2\right] \tag{8}$$

$$P_t^T\left(VIX_t, K\right) = e^{-r(T-t)}\left[K \cdot (1 - \Pi_2) - F_t^T \cdot (1 - \Pi_1)\right] \tag{9}$$

where r is the riskless interest rate and F_t^T is VIX future price given in (7). The two probabilities Π_1 and Π_2 are expressed as follows:

$$\begin{aligned}\Pi_1 &= \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left\{\frac{g\left(X_t, \tau; s-i\right) e^{-is\log K}}{g\left(X_t, \tau; -i\right) is}\right\} ds \\ \Pi_2 &= \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left\{\frac{g\left(X_t, \tau; s\right) e^{-is\log K}}{is}\right\} ds\end{aligned} \tag{10}$$

Proof. See Appendix B. □

As the integrands in (10) are singular at the required evaluation point $s = 0$, the Fast Fourier Transform (FFT) is unable to be used to calculate the above integrals. Therefore, we adopt the direct integration approach, which is sufficiently accurate for our purpose. Moreover, [7] showed that an efficient implementation of the direct integration method results in a sizable speed up of the calibration of stochastic volatility models. In this paper, we apply Gaussian quadrature to approximately calculate the above integrals. Gauss−Legendre quadrature is one of the most widely used numerical integration methods with high precision and stability. The abscissae in the Gaussian quadrature function are defined as the roots of the Legendre polynomial. Fortunately, the abscissae and their corresponding weights have been extensively tabulated. Thanks to pre-calculated abscissae and weights, numerical calculation can be easily implemented to solve the above integrals.

§4 Calibration procedure

In order to calibrate model parameters and investigate the performance of different models on pricing VIX options in the following section, we use the filtered dataset as stated in Section 2. The yield on three-month Treasury bills is the proxy for the riskless interest rate. In the calibration procedure, model parameters are estimated by minimizing a loss function. Therefore, the choice of loss functions is crucial. Among a variety of loss functions, the most commonly used one to calibrate parameters in option valuation is the mean square error (MSE) function which measures the deviation of market prices and model prices. However, [23] indicated that the MSE function assigns more weight to ITM option contracts and less weight to OTM options. In fact, investors are more concerned with OTM options especially when market prospects are uncertain. Therefore, we employ the mean logarithmic square error (MLSE) as the loss function in the model calibration procedure, defined as:

$$\mathrm{MLSE} = \frac{1}{N_1 N_2}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\left(\log C_{i,j} - \log C_{i,j}(\Theta)\right)^2 \tag{11}$$

with N_1 and N_2 respectively denoting the number of maturities and the number of strikes for each fixed maturity, where $C_{i,j}$ is the mid-prices of VIX call options and $C_{i,j}(\Theta)$ is the model-determined prices for a given parameter set Θ.

Table 4: VIX parameter estimates.

	Panel A: Single-factor volatility models		***Panel B: Multi-factor volatility models***		
	SSV	SSV-UJ	MSV	MSV-UJ	MSV-AJ
k	3.0648	3.2501	3.4531	3.3169	3.3289
	(0.0983)	(0.0911)	(0.0899)	(0.0886)	(0.1004)
θ	2.9192	2.5582	2.8271	2.4263	2.4971
	(0.0571)	(0.0309)	(0.0189)	(0.0196)	(0.0224)
k_1	8.0532	7.7479	5.1332	4.4058	4.1880
	(0.1634)	(0.1658)	(0.1080)	(0.0947)	(0.1409)
θ_1	0.7062	0.6909	0.6895	0.4152	0.5038
	(0.0331)	(0.0465)	(0.0257)	(0.0380)	(0.0421)
σ_1	2.5581	1.4466	2.7888	2.0274	1.8436
	(0.0724)	(0.1046)	(0.0838)	(0.0901)	(0.0974)
ρ_1	1.000	0.8135	0.9545	0.8416	0.8379
	(0.0000)	(0.0071)	(0.0099)	(0.0061)	(0.0068)
V_{10}	0.9506	0.5436	0.2692	0.2496	0.2192
	(0.0245)	(0.0581)	(0.0133)	(0.0192)	(0.0202)
k_2	-	-	11.9406	9.6995	10.0215
	-	-	(0.1202)	(0.1310)	(0.1928)
θ_2	-	-	0.2647	0.3495	0.3317
	-	-	(0.0388)	(0.0392)	(0.0457)
σ_2	-	-	5.6254	2.3773	2.5894
	-	-	(0.0735)	(0.0956)	(0.1367)
ρ_2	-	-	0.7316	0.6418	0.6583
	-	-	(0.0126)	(0.0076)	(0.0101)
V_{20}	-	-	1.2814	0.4626	0.5227
	-	-	(0.0275)	(0.0390)	(0.0433)
λ	-	2.9205	-	3.3359	3.9826
	-	(0.2681)	-	(0.1866)	(0.2177)
$1/\eta_1$	-	0.2206	-	0.2778	0.2890
	-	(0.0913)	-	(0.0659)	(0.0710)
$1/\eta_2$	-	-	-	-	0.1892
	-	-	-	-	(0.0495)
p	-	-	-	-	0.7263
	-	-	-	-	(0.0155)

This table shows the parameter calibration results for the single-factor volatility models (Panel A) and the multi-factor volatility models (Panel B) using VIX option filtered data during the sample period. The parameters are calibrated from the daily updated frequency by minimizing the loss function, defined as MLSE $= \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} (\log C_{i,j} - \log C_{i,j}(\Theta))^2$. For each parameter, we report the mean level and the standard error (in parentheses).

The parameter calibration results are reported in Table 4. There are several observations in order. First, the correlations between spot VIX and instantaneous variance are found to be positive, which means all models are capable of producing upward volatility skew. Second, within the multi-factor model framework, one volatility factor has low mean-reversion speed, k_1, which captures long-term fluctuations of volatility, while another with high mean-reversion speed, k_2, accounts for short-term fluctuations of volatility, following the discussions of [21]. Third, the models with jumps yield very different variance dynamics than the models without jumps. Specifically, the introduction of jump components significantly reduces almost all parameter values of the stochastic variance processes, whether for single-factor models or multi-factor models. For example, the MSV-UJ model has lower volatilities of variance ($\sigma_1 = 2.03$, $\sigma_2 = 2.38$) than the MSV model ($\sigma_1 = 2.79$, $\sigma_2 = 5.63$). This is consistent with expectations because part of variations of VIX are accounted by jumps. In other words, adding jumps makes a large difference in the estimation associated with the variance dynamics of the VIX, which in turn may have an impact on the pricing of VIX options. Finally, focusing on the parameters of the MSV-AJ model, we find that the probability of upward jumps is about 0.73 and the mean size of upward jumps ($1/\eta_1$=0.29) is higher than that of downward jumps ($1/\eta_2$=0.19). These observations are in line with the empirical analyses in Section 2.

§5 Empirical results

5.1 Pricing performance

To ensure the robustness of our results, we consider three measurements of the pricing errors, the mean absolute error (MAE), the root mean-squared error (RMSE) and the mean absolute percentage error (MAPE), defined as:

$$\mathrm{MAE} = \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} |C_{i,j} - C_{i,j}(\Theta)| \tag{12}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} (C_{i,j} - C_{i,j}(\Theta))^2} \tag{13}$$

$$\mathrm{MAPE} = \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \frac{|C_{i,j} - C_{i,j}(\Theta)|}{C_{i,j}} \tag{14}$$

In additional to the in-sample pricing performance, we also calculate the out-of-sample pricing errors. Following [8], we use the in-sample estimated parameters to compute option prices for the next day. The calculated results for each model are presented in Table 5. In general, more sophisticated models, the MSV-type models, significantly reduce in-sample and out-of-sample pricing errors regardless of the error functions. For example, the in-sample pricing

Table 5: Pricing errors for each model.

Panel A: In-sample results

	SSV	SSV-UJ	MSV	MSV-UJ	MSV-AJ
MAE	0.1045	0.0951	0.0861	0.0776	0.0749
RMSE	0.1627	0.1503	0.1386	0.1309	0.1261
MAPE	10.78	9.31	7.82	6.63	6.45

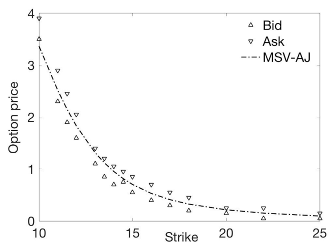
Panel B: Out-of-sample results

	SSV	SSV-UJ	MSV	MSV-UJ	MSV-AJ
MAE	0.1381	0.1240	0.1162	0.1053	0.1019
RMSE	0.2113	0.1947	0.1785	0.1694	0.1636
MAPE	13.43	12.28	10.03	8.72	8.56

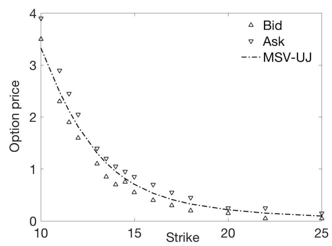
This table reports the average mean absolute errors (MAE), the average root mean-squared errors (RMSE), the average mean absolute percentage errors (MAPE) for each model during the sample period, with Panel A and B respectively presenting the in-sample and out-of-sample results. MAPE are reported in percentage.

errors of the MSV model measured by MAE, RMSE and MAPE are 0.0861, 0.1386 and 7.82%, respectively, while the in-sample pricing errors of the SSV model measured by MAE, RMSE and MAPE are 0.1045, 0.1627 and 10.78%, respectively. In other words, models with the second volatility factor generally outperform models without the second volatility factor. The consistency of both the in-sample and out-of-sample results suggests that the extra parameters do not cause over-fitting. In order to observe the difference of explaining option prices across alternative models more intuitively, Figure 3 depicts model-determined option prices along with market quotes as a function of strikes under time-to-maturity 21 days. The figure shows that the prices of the multi-factor models basically lie within the bid-ask band of quotes, while other models perform worse in describing the option quotes, especially for OTM option contracts. These findings provide support for the use of multi-factor volatility models as the means of capturing stochastic skew.

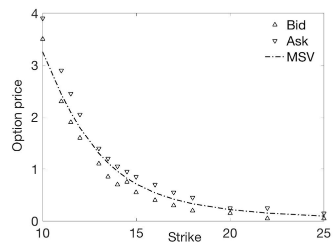
To further dissect the pricing errors across different moneyness and time-to-maturity, the results are classified into three moneyness groups: ITM options with $m < -0.1$, near-the-money (NTM) options with $-0.1 \leq m \leq 0.1$ and OTM options with $m > 0.1$, and each group includes three types of time-to-maturity as follows: short-term (less than 60 days), intermediate-term (between 60 and 120 days) and long-term (over 120 days). For each category, we report the in-sample and out-of-sample MAE, RMSE and MAPE of different models in Table 6 and 7. We observe that the MSV-type models outperform the SSV-type models across almost all



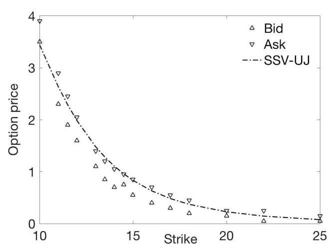
(a) MSV-AJ



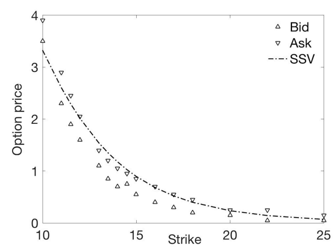
(b) MSV-UJ



(c) MSV



(d) SSV-UJ



(e) SSV

Figure 3: Comparison between VIX option quotes and model-determined prices.

moneyness and time-to-maturity categories. Among the MSV-type models, the MSV models with jumps perform better for most categories, especially for OTM options. Furthermore, we find from Table 5 that allowing for downward jumps reduces option pricing errors to a certain extent by comparing between the MSV-UJ model and the MSV-AJ model. Since the OTM option contracts tend to be more popular in VIX options markets, the accurate pricing for this type of option is more crucial. These findings confirm the importance of the multi-factor structure and asymmetric jumps for the valuation of VIX options.

As discussed in Section 2, the two volatility components are specified as the means of capturing stochastic skew, and the models with only one factor fail to capture the common feature. Accordingly, in order to illustrate the value of the multi-factor specification in capturing the shape of the implied volatility skew, Figure 4 depicts the model-determined implied volatility skews along with the market implied volatility skew extracted from OTM options with time-to-maturity 21 days. From the figure, the MSV-type models exhibits superior performance in fitting the implied volatility skew, although all the models are capable of displaying the upward volatility skew pattern. The result provides support for the valuable role of the multi-factor volatility structure in fitting the market implied volatility skew.

5.2 Pricing improvements

Next, we aim to analyze the marginal contributions of the additional components in explaining the VIX option prices. To do this, we make comparative studies across alternative

Table 6: Pricing errors across moneyness and time-to-maturity: In-sample

	MAE			RMSE			MAPE		
	Short	Intermediate	Long	Short	Intermediate	Long	Short	Intermediate	Long
ITM options									
SSV	0.3526	0.2183	0.5002	0.4592	0.2583	0.6150	13.08	5.53	9.14
SSV-UJ	0.3312	0.1940	0.4905	0.4419	0.2405	0.6097	11.62	4.29	9.26
MSV	0.3269	0.1952	0.4810	0.4381	0.2420	0.6012	11.68	4.35	8.51
MSV-UJ	0.3261	0.1909	0.4773	0.4357	0.2369	0.5982	11.43	4.24	8.53
MSV-AJ	0.3142	0.1928	0.4738	0.4253	0.2392	0.5930	11.07	4.23	8.44
NTM options									
SSV	0.1976	0.1593	0.2462	0.2668	0.2011	0.3028	12.67	5.07	6.22
SSV-UJ	0.1849	0.1482	0.2210	0.2474	0.1862	0.2705	10.23	5.01	5.30
MSV	0.1557	0.1462	0.2141	0.2063	0.1781	0.2594	8.52	4.32	4.91
MSV-UJ	0.1483	0.1378	0.2059	0.2017	0.1746	0.2546	8.05	4.20	4.60
MSV-AJ	0.1416	0.1377	0.1901	0.1929	0.1744	0.2356	7.71	4.17	4.27
OTM options									
SSV	0.0635	0.0741	0.0839	0.0967	0.1082	0.1194	13.88	8.76	7.92
SSV-UJ	0.0536	0.0701	0.0754	0.0851	0.0994	0.1073	12.52	6.89	6.76
MSV	0.0467	0.0613	0.0695	0.0782	0.0906	0.0985	10.07	6.02	5.89
MSV-UJ	0.0383	0.0510	0.0611	0.0656	0.0837	0.0906	8.11	5.14	4.93
MSV-AJ	0.0377	0.0494	0.0572	0.0648	0.0812	0.0825	7.88	5.05	4.79

The in-sample results are sorted by three moneyness levels and three time-to-maturity categories: ITM options with $m < -0.1$, NTM options with $-0.1 \leq m \leq 0.1$ and OTM options with $m > 0.1$; short-term (less than 60 days), intermediate-term ($60-120$ days) and long-term (over 120 days).

option pricing models with or without the additional components such as the second volatility factor or jumps. And we focus particularly on three comparisons. First, we compare the models without the second volatility factor with the models with the second volatility factor to investigate the marginal contributions of the second volatility factor to the valuation of VIX options. The second comparison is made between the MSV model and the MSV-UJ model to observe the contributions made by jump components in the context of the multi-factor volatility framework, which answers the question of whether the second volatility factor eliminates the need for jumps. Third, we investigate the marginal contributions of incorporating downward jumps within the multi-factor model framework by comparing the models without asymmetric jumps with the MSV-AJ model.

Following [26], we adopt the measurement, $\Delta\mathrm{RMSE}_{i|j}$, to measure the improvements in the

Table 7: Pricing errors across moneyness and time-to-maturity: Out-of-sample

	MAE			RMSE			MAPE		
	Short	Intermediate	Long	Short	Intermediate	Long	Short	Intermediate	Long
ITM options									
SSV	0.4779	0.4252	0.5323	0.6235	0.5404	0.6867	15.12	10.14	9.96
SSV-UJ	0.4713	0.4100	0.5298	0.6110	0.5375	0.6834	15.01	9.98	10.05
MSV	0.4592	0.3860	0.5109	0.5934	0.5012	0.6541	14.92	8.97	9.59
MSV-UJ	0.4416	0.3772	0.5168	0.5907	0.4932	0.6579	13.89	8.64	9.65
MSV-AJ	0.4279	0.3620	0.5041	0.5711	0.4740	0.6497	13.51	8.29	9.41
NTM options									
SSV	0.2264	0.2729	0.2983	0.2879	0.3092	0.3620	12.92	7.98	7.03
SSV-UJ	0.2142	0.2663	0.2861	0.2785	0.3004	0.3593	12.69	8.25	7.18
MSV	0.2054	0.2595	0.2787	0.2578	0.2891	0.3369	11.60	7.51	6.36
MSV-UJ	0.1943	0.2411	0.2792	0.2524	0.2872	0.3344	10.84	7.38	6.47
MSV-AJ	0.1876	0.2335	0.2582	0.2431	0.2784	0.3178	10.55	7.15	6.03
OTM options									
SSV	0.0893	0.0993	0.0975	0.1381	0.1443	0.1356	19.78	10.06	7.93
SSV-UJ	0.0714	0.0880	0.0835	0.1191	0.1204	0.1228	17.82	8.94	7.10
MSV	0.0583	0.0819	0.0865	0.0904	0.1157	0.1236	12.91	7.81	7.08
MSV-UJ	0.0478	0.0706	0.0764	0.0768	0.1082	0.1122	10.56	7.16	6.14
MSV-AJ	0.0475	0.0689	0.0723	0.0762	0.1049	0.1063	10.48	7.02	5.94

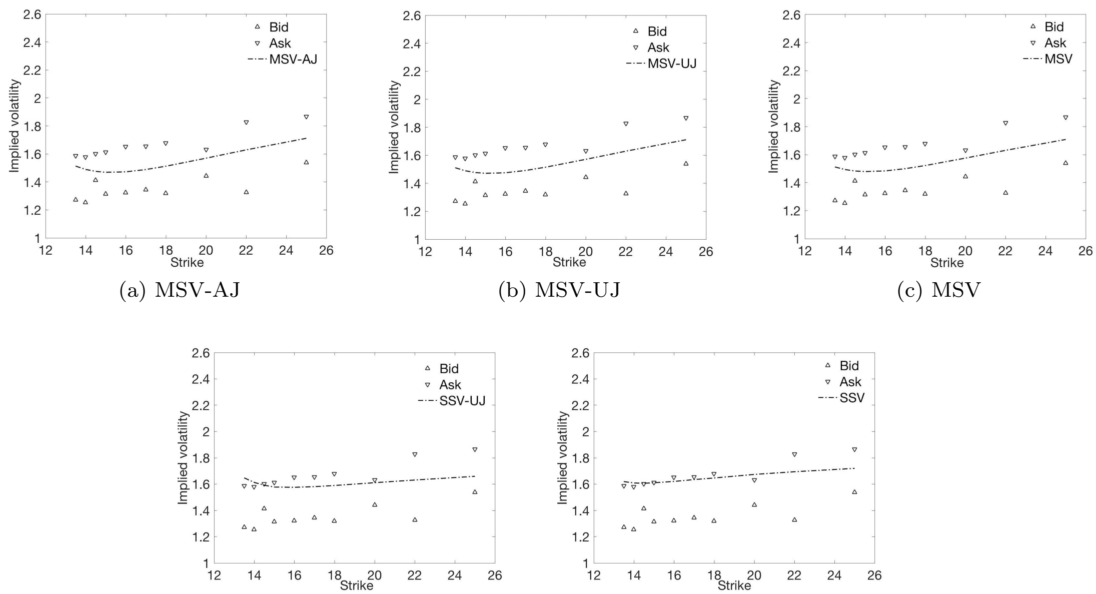
The out-of-sample results are sorted by three moneyness levels and three time-to-maturity categories: ITM options with $m < -0.1$, NTM options with $-0.1 \leq m \leq 0.1$ and OTM options with $m > 0.1$; short-term (less than 60 days), intermediate-term (60−120 days) and long-term (over 120 days).

pricing of VIX options made by the additional components. $\Delta\text{RMSE}_{i|j}$ is defined as:

$$\Delta\text{RMSE}_{i|j} = 100 \times (\log \text{RMSE}_i - \log \text{RMSE}_j), \tag{15}$$

where RMSE_i and RMSE_j respectively denote the RMSE of the model i and the RMSE of the model j. Table 8 shows in-sample and out-of-sample pricing improvements across different models. $\Delta\text{RMSE}_{i|j}$ represents the improvements in the pricing of VIX options made by a model i relative to a model j. For example, $\Delta\text{RMSE}_{\text{MSV-UJ}|\text{MSV}}$ describes the marginal contributions for jump components when the second volatility factor are specified.

The main results of Tables 8 are as follows. First, within the single-factor volatility model framework, the in-sample and out-of-sample pricing improvements attributable to jumps are 7.93% and 8.18%, respectively. And compared with the MSV model, the MSV-UJ model greatly



(a) MSV-AJ (b) MSV-UJ (c) MSV

(d) SSV-UJ (e) SSV

Figure 4: The model implied volatility skews and the corresponding market values.

improves the valuation of options by 5.72% in the in-sample tests and by 5.23% in the out-of-sample tests. These empirical findings consistently identify the value of jumps in the pricing of options, whether within the single-factor volatility model framework or within the multi-factor framework, and illustrate that the need for jumps cannot be eliminated by the second volatility factor.

Second, a comparison of the SSV-UJ model and the MSV-UJ model reveals that the pricing improvements made by the second volatility factor are small when jumps are already accounted for, with values 13.82% for in-sample tests and 13.92% for out-of-sample tests. However, the improvements made by the second volatility factor are relative large when jumps are not incorporated, with values 16.03% for in-sample tests and 16.87% for out-of-sample tests. The above phenomenon is due to the reason that jumps have contributed a little to explaining option prices.

Third, benchmarking against the SSV model, the pricing improvements provided by the inclusion of jumps are lower than those provided by the second volatility factor. Specifically, the in-sample and out-of-sample pricing improvements made by jumps are 7.93% and 8.18%, respectively, while those made by the second volatility factor are 16.03% and 16.87%, respectively. That is to say, the second volatility factor contributes more than jumps in improving the explanatory power on option prices.

Lastly, the pricing improvements attributable to downward jumps relative to the MSV-UJ model are positive in both in-sample and out-of-sample tests, although the magnitude of values is small, with 3.74% and 3.48%, respectively, suggesting that the MSV-AJ model slightly outperforms the MSV-UJ model in describing the option prices. To summarize, we make comparative studies on the option pricing performance across five alternative models to investigate whether incorporating an additional volatility factor and asymmetric jumps is significant for valuation of VIX options, and to further indicate the importance of capturing stochastic skew and asymmetric jumps features. The above results provide solid evidence that inclusion of the second volatility factor and asymmetric jumps indeed raises the precision of the valuation of VIX options, due to providing more flexibility in the modeling of VIX dynamics. And the need for jumps cannot be eliminated by the multi-factor volatility structure.

Table 8: Pricing improvements across different models.

Panel A: In-sample tests

	$\Delta\text{RMSE}_{i\|\text{SSV}}$	$\Delta\text{RMSE}_{i\|\text{SSV-UJ}}$	$\Delta\text{RMSE}_{i\|\text{MSV}}$	$\Delta\text{RMSE}_{i\|\text{MSV-UJ}}$
SSV	0.00	7.93	16.03	21.75
SSV-UJ	-7.93	0.00	8.10	13.82
MSV	-16.03	-8.10	0.00	5.72
MSV-UJ	-21.75	-13.82	-5.72	0.00
MSV-AJ	-25.48	-17.56	-9.45	-3.74

Panel B: Out-of-sample tests

	$\Delta\text{RMSE}_{i\|\text{SSV}}$	$\Delta\text{RMSE}_{i\|\text{SSV-UJ}}$	$\Delta\text{RMSE}_{i\|\text{MSV}}$	$\Delta\text{RMSE}_{i\|\text{MSV-UJ}}$
SSV	0.00	8.18	16.87	22.10
SSV-UJ	-8.18	0.00	8.69	13.92
MSV	-16.87	-8.69	0.00	5.23
MSV-UJ	-22.10	-13.92	-5.23	0.00
MSV-AJ	-25.59	-17.40	-8.72	-3.48

This table reports the in-sample and out-of-sample pricing improvements across different models. $\Delta\text{RMSE}_{i|j}$ measures the improvements in the pricing of VIX options made by a model i relative to a model j, defined as $\Delta\text{RMSE}_{i|j} = 100 \times (\log \text{RMSE}_i - \log \text{RMSE}_j)$. A negative (positive) value of $\Delta\text{RMSE}_{i|j}$ indicates that the model i outperforms (underperforms) the model j.

§6 Conclusion

Given these distinctive features observed in VIX markets, such as upward implied volatility skew, stochastic skew and asymmetric jumps, this paper develops a general valuation model incorporating these features and investigates their marginal contributions to the valuation of VIX options by making comparative studies across different models. Within the model framework, we derive quasi-analytical solutions for the price of VIX derivatives.

The main conclusions are drawn from the empirical results as follows. First, both in-sample and out-of-sample pricing performance significantly favor the MSV-type models over the SSV-type models, especially for OTM options. Besides, in terms of fitting the implied volatility skew, the MSV-type models exhibits superior performance, due to the capability of capturing stochastic skew. The empirical results consistently suggest that the introduction of the second volatility factor has significant implications on explaining the option prices. Second, the value of jumps is identified in the valuation of options, even after specifying the second volatility factor. In addition, the consistent findings indicate that allowing for downward jumps can improve the VIX option pricing to a certain extent. Finally, we investigate the relative value of an additional volatility factor and jumps in VIX option pricing. Comparisons of the pricing improvements made by an additional volatility factor and those made by jumps reveal that the second volatility factor contributes more than jumps in improving the explanatory power on option prices. In summary, two volatility factors and jump components play important roles in the pricing of VIX options. The multi-factor stochastic volatility model with asymmetric jumps, which adequately capture all of the features in VIX markets, performs better than other alternative models in fitting VIX option quotes, both in-sample and out-of-sample.

Appendix

Appendix A. The conditional characteristic function

For notational simplicity, we denote the conditional characteristic function (4) by $g{=}g\left(X_t,\tau;s\right)$. Given the model specification (2), the differential of g can be computed by applying Itô's lemma as follows:

$$\begin{aligned}dg =&\left[\frac{\partial g}{\partial t}+k(\theta-X_t)\frac{\partial g}{\partial X_t}+\sum_{i=1}^{2}k_i(\theta_i-V_{it})\frac{\partial g}{\partial V_{it}}+\frac{1}{2}\sum_{i=1}^{2}V_{it}\frac{\partial^2 g}{\partial X_t^2}+\frac{1}{2}\sum_{i=1}^{2}\sigma_i^2V_{it}\frac{\partial^2 g}{\partial V_{it^2}}\right.\\&\left.+\sum_{i=1}^{2}\sigma_i\rho_iV_{it}\frac{\partial^2 g}{\partial X_t\partial V_{it}}\right]dt+\left[g(X_t+Y,\tau;s)-g(X_t,\tau;s)\right]dN_t\\&+\sqrt{V_{1t}}\frac{\partial g}{\partial X_t}dW_t^1+\sigma_1\sqrt{V_{1t}}\frac{\partial g}{\partial V_{1t}}\left(\rho_1dW_t^1+\sqrt{1-\rho_1^2}dW_t^3\right)\\&+\sqrt{V_{2t}}\frac{\partial g}{\partial X_t}dW_t^2+\sigma_2\sqrt{V_{2t}}\frac{\partial g}{\partial V_{2t}}\left(\rho_2dW_t^2+\sqrt{1-\rho_2^2}dW_t^4\right)\end{aligned}\tag{A.1}$$

By iterated conditioning argument, it is easy to prove that g is a Q-martingale. Consequently, taking conditional expectations on both sides of the above equation obtains:

$$\begin{aligned}0 =&\frac{\partial g}{\partial t} + k(\theta - X_t)\frac{\partial g}{\partial X_t} + \sum_{i=1}^{2} k_i(\theta_i - V_{it})\frac{\partial g}{\partial V_{it}} + \frac{1}{2}\sum_{i=1}^{2} V_{it}\frac{\partial^2 g}{\partial X_t^2} + \frac{1}{2}\sum_{i=1}^{2}\sigma_i^2 V_{it}\frac{\partial^2 g}{\partial V_{it}^2}\\ &+ \sum_{i=1}^{2}\sigma_i\rho_i V_{it}\frac{\partial^2 g}{\partial X_t \partial V_{it}} + \lambda E_t^{\mathbb{Q}}[g(X_t + Y, \tau; s) - g(X_t, \tau; s)]\end{aligned} \tag{A.2}$$

Because of the affine structure of the model, we surmise that g has the following exponential affine form:

$$g\left(X_t, \tau; s\right) = exp[A(\tau; s) + \sum_{i=1}^{2} B_i(\tau; s)V_{it} + C(\tau; s)X_t] \tag{A.3}$$

with initial conditions $A\left(0; s\right) = 0, B_i\left(0; s\right) = 0(i = 1, 2)$ and $C\left(0; s\right) = is$. Given the probability density function of y, substituting this form into partial differential equation (A.2) gives:

$$\begin{aligned}0 =& -\left(\frac{\partial A}{\partial \tau} + \sum_{i=1}^{2} V_{it}\frac{\partial B_i}{\partial \tau} + X_t\frac{\partial C}{\partial \tau}\right) + k(\theta - X_t)C + \sum_{i=1}^{2} k_i(\theta_i - V_{it})B_i + \frac{1}{2}\sum_{i=1}^{2} V_{it}C^2\\ &+ \frac{1}{2}\sum_{i=1}^{2}\sigma_i^2 V_{it}B_i^2 + \sum_{i=1}^{2}\sigma_i\rho_i V_{it}B_i C + \lambda\left(\frac{p\eta_1}{\eta_1 - C} + \frac{q\eta_2}{\eta_2 + C} - 1\right)\end{aligned} \tag{A.4}$$

By the arbitrariness of V_{it} and X_t, we have the following system of ODEs:

$$\begin{aligned}\frac{\partial A}{\partial \tau} =& k\theta C + \sum_{i=1}^{2} k_i\theta_i B_i + \lambda\left(\frac{p\eta_1}{\eta_1 - C} + \frac{q\eta_2}{\eta_2 + C} - 1\right)\\ \frac{\partial B_i}{\partial \tau} =& \frac{1}{2}C^2 + (C\rho_i\sigma_i - k_i)B_i + \frac{1}{2}B_i^2\sigma_i^2 \qquad for\ i = 1, 2\\ \frac{\partial C}{\partial \tau} =& -kC\end{aligned} \tag{A.5}$$

It is easy to obtain the explicit formula of $C\left(\tau; s\right)$ which is given by $C\left(\tau; s\right) = ise^{-k\tau}$. Plugging $C\left(\tau; s\right)$ into the ODE for $B_i\left(\tau; s\right)$ yields the Riccati equations which can be expressed by Kummer functions via changes of variables (see, [1]). However, the calculation is not stable and rather time-consuming. Therefore, we recommend using Runge-Kutta numerical method to solve the ODE for $B_i\left(\tau; s\right)$. With $C\left(\tau; s\right)$ and $B_i\left(\tau; s\right)$ availabe, $A\left(\tau; s\right)$ can be easily solved by integrating both sides of (A.5).

Appendix B. The pricing formulas for VIX options

The price of a European call on the VIX at time t is expressed as discounted conditional expectation of terminal payoff under the $\mathbb{Q}$ measure as follows:

$$C_t^T\left(VIX_t, K\right) = e^{-r(T-t)}E_t^{\mathbb{Q}}[\max\left(VIX_T - K, 0\right)]$$

Given the characteristic function and the VIX future contracts pricing formula (7), by making the change of measure, VIX call option price can be further expressed as

$$\begin{aligned}
C_t^T(VIX_t, K) &= e^{-r(T-t)} E_t^{\mathbb{Q}}\Big(e^{X_T}\mathbb{I}_{\{X_T\geq \log K\}} - K\mathbb{I}_{\{X_T\geq \log K\}}\Big) \\
&= e^{-r(T-t)}\left\{E_t^{\mathbb{Q}}(e^{X_T})E_t^{\mathbb{Q}}\Big[\frac{e^{X_T}/E^{\mathbb{Q}}(e^{X_T})}{E_t^{\mathbb{Q}}[e^{X_T}/E^{\mathbb{Q}}(e^{X_T})]}\mathbb{I}_{\{X_T\geq \log K\}}\Big] - KE_t^{\mathbb{Q}}(\mathbb{I}_{\{X_T\geq \log K\}})\right\} \\
&= e^{-r(T-t)}\Big[E_t^{\mathbb{Q}}(e^{X_T})E_t^{\mathbb{Q}_1}(\mathbb{I}_{\{X_T\geq \log K\}}) - KE_t^{\mathbb{Q}}(\mathbb{I}_{\{X_T\geq \log K\}})\Big] \\
&:= e^{-r(T-t)}\left[F_t^T\cdot\Pi_1 - K\cdot\Pi_2\right] \qquad (A.6)
\end{aligned}$$

where the $\mathbb{Q}_1$ measure is defined by the following Esscher transform:

$$\left.\frac{d\mathbb{Q}_1}{d\mathbb{Q}}\right|_{\mathcal{F}_t} = \frac{e^{X_t}}{E^{\mathbb{Q}}\left[e^{X_t}\right]}$$

Under the $\mathbb{Q}_1$ measure, the conditional characteristic function of X_T is given by:

$$E_t^{\mathbb{Q}_1}\left[e^{isX_T}\right] = E_t^{\mathbb{Q}}\left[\frac{e^{X_T}}{E_t^{\mathbb{Q}}\left[e^{X_T}\right]}e^{isX_T}\right] = \frac{g\left(X_t,\tau;s-i\right)}{g\left(X_t,\tau;-i\right)}$$

With the conditional characteristic functions under the $\mathbb{Q}$ measure and $\mathbb{Q}_1$ measure available, the two probabilities in (10) can be solved by inversion theorem of [11]:

$$\begin{aligned}
\Pi_1 &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left\{\frac{g\left(X_t,\tau;s-i\right)e^{-is\log K}}{g\left(X_t,\tau;-i\right)is}\right\}ds \\
\Pi_2 &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left\{\frac{g\left(X_t,\tau;s\right)e^{-is\log K}}{is}\right\}ds
\end{aligned} \qquad (A.7)$$

We have derived the pricing formula (8) for VIX call options. The pricing formula (9) for put options can be derived similarly or obtained by using the modified put-call parity.

References

[1] A Polyanin, V Zaitsev. *Handbook of Exact Solutions for Ordinary Differential Equations*, USA: Chapman & Hall/CRC, 2003.

[2] A Sepp. *VIX Option Pricing in a Jump-Diffusion Model*, Risk Magazine, 2008, 21: 84-89.

[3] A Kaeck, C Alexander. *Continuous-time VIX dynamics: On the role of stochastic volatility of volatility*, International Review of Financial Analysis, 2013, 28(327): 46-56.

[4] C L Lo, P T Shih, Y H Wang, M T Yu. *VIX derivatives: Valuation models and empirical evidence*, Pacific-Basin Finance Journal, 2019, 53: 1-21.

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

[6] D Psychoyios, G Dotsis, Markellos R N. *A jump diffusion model for VIX volatility options and futures*, Review of Quantitative Finance and Accounting, 2010, 35(3): 245-269.

[7] F Kilin. *Accelerating the Calibration of Stochastic Volatility Models*, Journal of Derivatives, 2011, 18(3): 7-16.

[8] G Bakshi, C Cao, Z Chen. *Empirical performance of alternative option pricing models*, Journal of Finance, 1997, 52 (5): 2003-2049.

[9] G Dotsis, D Psychoyios, G Skiadopoulos. *An empirical comparison of continuous-time models of implied volatility indices*, Journal of Banking and Finance, 2007, 31(12): 3584-3603.

[10] G Drimus, W Farkas. *Local volatility of volatility for the VIX market*, Review of Derivatives Research, 2013, 16(3): 267-293.

[11] J Gil-Pelaez. *Note on the inversion theorem*, Biometrika, 1996, 38(4): 443-458.

[12] J Detemple, C Osakwe. *The valuation of volatility options*, Review of Finance, 2000, 4(1): 21-50.

[13] J Mencía, E Sentana. *Valuation of VIX derivatives*, Journal of Financial Economics, 2013, 108(2): 367-391.

[14] J Goard, M Mazur. *Stochastic volatility models and the pricing of VIX options*, Mathematical Finance, 2013, 23(3): 439-458.

[15] J M Romo. *Pricing volatility options under stochastic skew with application to the VIX index*, The European Journal of Finance, 2017, 23(4): 353-374.

[16] J Detemple, Y Kitapbayev. *On American VIX options under the generalized 3/2 and 1/2 models*, Mathematical Finance, 2018, 28(2): 550-581.

[17] M Grasselli, J M Romo. *Stochastic Skew and Target Volatility Options*, Journal of Futures Markets, 2016, 36(2): 174-193.

[18] N Wagner, A Szimayer. *Local and spillover shocks in implied market volatility: evidence for the U.S. and Germany*, Research in International Business and Finance, 2004, 18(3) 237-251.

[19] P Carr, L Wu. *Stochastic Skew in currency options*, Journal of Financial Economics 2007, 86(1): 213-247.

[20] P Christoffersen, K Jacobs, K Ornthanalai, Y Wang. *Option valuation with long-run and short-run volatility components*, Journal of Financial Economics, 2008, 90(3): 272-297.

[21] P Christoffersen, S L Heston, K Jacobs. *The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work So Well*, Management Science 2009, 55(12): 1914-1932.

[22] P Christoffersen, K Jacobs, K Mimouni. *Volatility dynamics for the S&P 500: evidence from realized volatility, daily returns, and option prices*, Review of Financial Studies, 2010, 23(8): 3141-3189.

[23] R Rebonato. *Volatility and Correlation: The Perfect Hedger and The Fox*, John Wiley and Sons, 2004.

[24] S G Kou. *A jump-diffusion model for option pricing*, Management Science, 2002, 48(8): 1086-1101.

[25] X Xu, S J Taylor. *The Term Structure of Volatility Implied by Foreign Exchange Options*, Journal of Financial and Quantitative Analysis, 1994, 29(1): 57-74.

[26] Y H Park. *The Effects of Asymmetric Volatility and Jumps on the Pricing of VIX Derivatives*, Journal of Econometrics, 2016, 192(1): 313-328.

Department of Mathematics, Zhejiang University, Hangzhou 310027, China.
Email: bojing@zju.edu.cn, shli@zju.edu.cn, xiaoyutan@zju.edu.cn