A Theory of the CBOE SKEW

Fang Zhen
Department of Accountancy and Finance
Otago Business School, University of Otago
Dunedin 9054, New Zealand
Email: fang.zhen@otago.ac.nz

Jin E. Zhang[1]
Department of Accountancy and Finance
Otago Business School, University of Otago
Dunedin 9054, New Zealand
Email: jin.zhang@otago.ac.nz

Keywords: The CBOE SKEW; Risk-neutral skewness; VIX; Jump-diffusion; Stochastic volatility

JEL Classification Code: G12; G13

[1]Corresponding author. Tel: 64 3 479 8575. Jin E. Zhang has been supported by an establishment grant from University of Otago.

A Theory of the CBOE SKEW

Abstract

The CBOE SKEW is a new index launched by the Chicago Board Options Exchange (CBOE) in February 2011. Its term structure tracks risk-neutral skewness of the S&P 500 (SPX) index for different maturities. In this paper, we develop a theory for the CBOE SKEW by modelling SPX using a jump-diffusion process with stochastic volatility and stochastic jump intensity. With the market data of term structures of VIX and SKEW, we estimate model parameters and obtain the four processes of instantaneous variance, jump intensity and their long-term mean levels. Our result can be used to describe SPX risk-neutral distribution and price SPX options.

1 Introduction

The CBOE SKEW is a new index launched by the Chicago Board Options Exchange (CBOE) in February 2011. Its term structure tracks risk-neutral skewness of the S&P 500 (SPX) index for different maturities. Skewness, which measures the asymmetry of a distribution, gives more precise details of the distribution of a underlying asset and improves the pricing accuracy of options written on this asset. Wang and Daigler (2012) examine the empirical relations between the SKEW, the VIX, the VIX of VIX and the SKEW of VIX. Faff and Liu (2014) propose an alternative measure of market asymmetry against the SKEW index. A theory of the CBOE SKEW has not been developed. In this paper, we establish a theory for the SKEW index by modelling SPX using a jump-diffusion process with stochastic volatility and stochastic jump intensity, and separate the diffusive variance and jump intensity using the market data of term structures of CBOE VIX and SKEW.

The traditional Capital Asset Pricing Model (CAPM) is mean-variance efficient while Harvey and Siddique (2000) show that it only holds if the pricing kernel is linear in the market return. Kraus and Litzenberger (1976) extend the traditional CAPM to the 3-moment CAPM to incorporate the effect of skewness preference and Mitton and Vorkink (2007) find evidence that investors deliberately choose under-diversified portfolios to increase skewness exposure. Empirical Studies identify the importance of skewness risk and investigate the pricing factors in the cross section of stock returns, such as the idiosyncratic skewness (see Boyer, Mitton and Vorkink (2010)) as opposed to the coskewness in 3-moment CAPM, the market skewness (see Chang, Christoffersen and Jacobs (2013)), or the ex ante risk-neutral skewness (see Conrad, Dittmar and Ghysels (2013)) extracted from option prices. Besides the effect of skewness to returns, the skewness itself is of great interest to researchers. The predictability power of trading volume and past returns to skewness is tested by Chen, Hong and Stein (2001), the cross-sectional variation in skewness is investigated by Dennis and Mayhew (2002), the investor sentiment effect to skewness is examined by Han (2008), and the stylized facts of positive firm-level skewness and negative aggregate skewness are reconciled by Albuquerque (2012).

Skewness is an indispensable risk factor in asset pricing studies, especially during financial crises. It also provides information for practitioners, such as portfolio management (see DeMiguel et al. (2013)). The CBOE launched SKEW index to track the S&P 500 tail risk in February 2011 using Bakshi, Kapadia and Madan's (2003) model-free methodology. Skewness is not tradable but speculators or hedgers can take a position in it by forming stock option portfolios (see Bali and Murray (2013)).

Skewness is the perceived asymmetry in stock returns, which can be explained by leverage-effect, volatility-feedback or bubble theories. All these theories are consistent in jump diffusion models, as leverage effect and volatility feedback explain the negative correlation between stock returns and Brownian variance innovations while bubbles represent the jumps in the stock prices. Intuitively, the heterogeneous believes reflected by turnovers boosts volatility, whereas the momentum reflected by positive past returns or bearishness increases jump intensity. Additionally, the Black-Scholes (1973) implied volatility smile/smirk and the down-to-up variance, used by Chen, Hong and Stein (2001) and Faff and Liu (2014), are alternative measures for the return asymmetry.

In this paper, we provide analytical formulas for the CBOE SKEW in various affine jump diffusion (AJD) models, decompose the market skewness into jump component and variance correlation component in a 5-factor model with the term structure data of CBOE VIX and SKEW. The skewness calculation formulas provide a linkage between skewness and jumps/variance-correlation as well as other fixed model parameters, which are crucial factors for option pricing.

A large variety of option pricing formulas are based on affine jump diffusion models due to their tractability. Heston (1993) develops a stochastic volatility model that allows arbitrary correlation between volatility and spot-asset returns and explains the return skewness which can not be captured by Black and Scholes (1973) model. However, Heston model constitutes an inadequate explanation of the instantaneous skewness (see Das and Sundaram (1999) and Zhang et al. (2015)). Merton (1976) constructs a jump-diffusion model where the jumps in returns could be another source for skewness. Bates (2000) examines the combined stochastic-volatility/jump-diffusion model for the negatively skewed post-crash returns. Eraker, Johannes and Polson (2003) test the AJD models with jumps in volatility which persistently steepen the slope of implied volatility curves in addition to the transient impact of jumps in returns. Overall, skewness is a nested outcome of variance correlation and jumps.

The instantaneous Brownian variance and jump intensity are important for asset pricing. However, these two factors are unobservable in financial markets. A strand of empirical studies exploring the return-skewness relationship use historical skewness proxies, which require further assumptions to be valid, such as ergodicity. Using the easily available VIX and SKEW term structure data, which are daily updated on the CBOE website, our paper provides an estimation procedure to quantify the risk-neutral Brownian variance and jump intensity, which could be used as more reliable risk factors to explain returns than historical proxies.

The jump risk and correlation risk are two different common risk factors for the explanation of returns, as jump risk mainly accounts for short-term skewness whereas correlation risk primarily affects long-term skewness. The physical version of these two factors can be captured by different constructions: the former through the daily third moment of log-returns and the latter through the correlation between daily log-return and future variance as in Neuberger (2012). Thus, extracting the risk neutral jump intensity and Brownian variance from VIX and SKEW term structure data, our paper provides a theoretical platform to quantify the jump risk premium and correlation risk premium.

The remainder of this paper is organized as follows. Section 2 introduces the CBOE SKEW index. Section 3 derives the formulas for CBOE SKEW in various jump diffusion models. Section 4 discusses the data selection. Section 5 presents the estimation procedure and analyzes the empirical performance. Section 6 concludes with suggestions for future research.

2 Definition of the CBOE SKEW

The CBOE SKEW is a new index launched by the Chicago Board Options Exchange (CBOE) on February 23, 2011. Its term structure tracks risk-neutral skewness of the S&P 500 (SPX) index for different maturities. The SKEW is computed from all of the out-of-the-money (OTM) SPX option prices by using Bakshi, Kapadia and Madan's (2003) methodology.

At time t, the SKEW is defined as

$$SKEW_t = 100 - 10 \times Sk_t, \tag{1}$$

where Sk_t is risk-neutral skewness given by

$$Sk_t = E_t^Q \left[\frac{(R_t^T - \mu)^3}{\sigma^3} \right]; \tag{2}$$

R_t^T is the logarithmic return of SPX at time T, denoted as S_T, against current forward price with maturity T, F_t^T; μ is the expected return and σ^2 is the variance in risk-neutral measure Q

$$R_t^T = \ln \frac{S_T}{F_t^T}, \qquad \mu = E_t^Q(R_t^T), \qquad \sigma^2 = E_t^Q[(R_t^T - \mu)^2]. \tag{3}$$

Expanding equation (2) gives

$$Sk_t = \frac{E_t^Q[(R_t^T)^3] - 3E_t^Q(R_t^T)E_t^Q[(R_t^T)^2] + 2[E_t^Q(R_t^T)]^3}{E_t^Q[(R_t^T)^2] - [E_t^Q(R_t^T)]^2}. \tag{4}$$

Following Bakshi, Kapadia and Madan (2003), the CBOE evaluates the first three moments of log return $E_t^Q(R_t^T)$, $E_t^Q[(R_t^T)^2]$ and $E_t^Q[(R_t^T)^3]$ in risk-neutral measure by using current prices of European options with all available strikes as follows

$$\begin{aligned}
E_t^Q(R_t^T) &= -e^{r\tau}\sum_i \frac{1}{K_i^2}Q(K_i)\Delta K_i + \ln\frac{K_0}{F_t^T} + \frac{F_t^T}{K_0} - 1, && (5)\\
E_t^Q[(R_t^T)^2] &= e^{r\tau}\sum_i \frac{2 - 2\ln\frac{K_i}{F_t^T}}{K_i^2}Q(K_i)\Delta K_i + \ln^2\frac{K_0}{F_t^T} + 2\ln\frac{K_0}{F_t^T}\left(\frac{F_t^T}{K_0} - 1\right), && (6)\\
E_t^Q[(R_t^T)^3] &= e^{r\tau}\sum_i \frac{6\ln\frac{K_i}{F_t^T} - 3\ln^2\frac{K_i}{F_t^T}}{K_i^2}Q(K_i)\Delta K_i + \ln^3\frac{K_0}{F_t^T}\\
&\quad +3\ln^2\frac{K_0}{F_t^T}\left(\frac{F_t^T}{K_0} - 1\right), && (7)
\end{aligned}$$

where F_t^T is the forward index level derived from SPX option prices using put-call parity; K_0 is the first listed price below F_t^T; K_i is the strike price of the ith OTM option (a call if $K_i > K_0$ and a put if $K_i < K_0$); ΔK_i is half the difference between strikes on either side of K_i, i.e., $\Delta K_i = \frac{1}{2}(K_{i+1} - K_{i-1})$, and for minimum (maximum) strike, ΔK_i is simply the distance to the next strike above (below); r is the risk-free interest rate; $Q(K_i)$ is the midpoint of bid-ask spread for each option with strike K_i; τ is the time to expiration as a fraction of a year. The reasoning behind equations (5), (6) and (7) are included in Appendix A.

The SKEW index refers to 30-day maturity, i.e., $\tau = \tau_0 \equiv 30/365$. In general, 30-day options are not available. The current 30-day skewness Sk_t is derived by inter- or extrapolation from the current risk-neutral skewness at adjacent expirations, Sk_t^{near} and Sk_t^{next} as follows

$$Sk_t = \omega Sk_{t,\text{near}} + (1-\omega)Sk_{t,\text{next}}, \tag{8}$$

where ω is a weight determined by

$$\omega = \frac{\tau_{\text{next}} - \tau_0}{\tau_{\text{next}} - \tau_{\text{near}}},$$

and τ_{near} and τ_{next} are the times to expiration (up to minute) of the near- and next-term options respectively. The near- and next-term options are usually the first and second SPX contract months. "Near-term" options must have at least one week to expiration in order to minimize possible close-to-expiration pricing anomalies. For near-term options with less than one week to expirations, the data rolls to the second SPX contract month. "Next-term" is the next contract month following near-term. While calculating time to expiration, the SPX options are deemed to expire at the open of trading on SPX settlement day, i.e., the third Friday of the month.

3 Model

Option-pricing models have been developed by using different kinds of stochastic processes for the underlying stock, such as jump-diffusion with stochastic volatility and stochastic jump intensity. The purpose of making volatility and jump intensity stochastic is to capture time-varying second and third moments of stock return. Traditionally these models are usually estimated by using some numerical approaches with volatility and jump intensity being latent variables. These numerical estimation procedures are often highly technical and very time-consuming.

With the observable information of two term structures of VIX and SKEW from the CBOE, it is now possible to estimate risk-neutral underlying process explicitly. The volatility and jump intensity processes are not latent any more. In fact, in this paper we will make them semi-observable. In order to achieve this goal, we need some theoretical results on the term structures of VIX and risk-neutral skewness implied in different models.

To intuit the result, we begin with the simplest case, i.e., the Black-Scholes (1973) model.

Proposition 1 *In the Black-Scholes (1973) model, the risk-neutral underlying stock is modeled by*

$$\frac{dS_t}{S_t} = rdt + \sigma dB_t,$$

where B_t is a standard Brownian motion, r is risk-free rate and σ is volatility. The model-implied squared VIX and skewness at time t for time to maturity τ are as follows

$$VIX_{t,\tau}^2 = \sigma^2, \qquad Sk_{t,\tau} = 0. \tag{9}$$

This is a trivial case. The log return is normally distributed, hence the volatility term structure is flat and the skewness is zero.

In order to create skewness, we need include jumps, e.g., Poisson process, into the model. Merton's (1976) jump-diffusion model is a pioneer along this direction.

Proposition 2 *In Merton (1976) model, the risk-neutral underlying stock is modeled by*

$$\frac{dS_t}{S_t} = rdt + \sigma dB_t + (e^x - 1)dN_t - \lambda(e^x - 1)dt,$$

where N_t is a poisson process with constant jump size x and jump intensity λ. The model-implied squared VIX and skewness at time t for time to maturity τ are as follows

$$VIX_{t,\tau}^2 = \sigma^2 + 2\lambda(e^x - 1 - x), \qquad Sk_{t,\tau} = \frac{\lambda x^3}{(\sigma^2 + \lambda x^2)^{3/2}}\ \tau^{-1/2}. \tag{10}$$

Proof. See Appendix B.1.

Remark 2.1. Here we present a result with a constant jump size in order to make the formulas of VIX and SKEW term structures simple and intuitive. It can be extended to an arbitrary distribution without much difficulty, see e.g., Zhang, Zhao and Chang (2012).

Remark 2.2. The VIX term structure is again flat with a squared VIX being $\sigma^2 + 2\lambda(e^x - 1 - x)$, which is different from annualized term variance (variance swap rate) $Var_{t,\tau} = \sigma^2 + \lambda x^2$. The difference is due to the CBOE definition of VIX. It is small for small x, see Luo and Zhang (2012) for a detailed discussion.

Remark 2.3. We notice that $\sqrt{\tau}\ Sk_{t,\tau}$ is a constant in this model. Hence we a simple criterion as follows: *If $\sqrt{\tau}\ Sk_{t,\tau}$ is not a constant, then Merton (1976) jump-diffusion model does not apply in the options market.*

Furthermore, we notice that $\sqrt{\tau}\ Sk_{t,\tau}$ is equal to the daily skewness if the daily returns are independent and identically distributed. Thus, we introduce a new concept of the regularized skewnss.

Definition The regularized skewness, $RSk_{t,\tau}$, at time t, is given by

$$RSk_{t,\tau} = \sqrt{\tau}\ Sk_{t,\tau} = \sqrt{\tau} E_t^Q \left[\frac{(R_{t,\tau} - \mu(R_{t,\tau}))^3}{\sigma^3(R_{t,\tau})} \right], \tag{11}$$

where $R_{t,\tau}$ denotes the logarithmic return over the period $[t, t+\tau]$, $\mu(R_{t,\tau})$ and $\sigma^2(R_{t,\tau})$ denote its mean and variance, respectively.

In addition to jumps, skewness can also be created by using stochastic volatility with leverage effect, i.e., correlation between stock return and volatility. Along this direction, Heston (1993) model has become a standard platform partially because of its analytical tractability due to an affine structure. Das and Sundaram (1999) first attempt to describe analytically the skewness implied in Heston model, unfortunately their claimed closed-form formula is not exact. Zhao, Zhang and Chang (2013) provide a partial result for a special case of zero mean-reverting speed, i.e., $\kappa = 0$. A full explicit formula is not available until the recent work of Zhang et al (2015).

Proposition 3 *(Zhang et al 2015) In the Heston (1993) model, the risk-neutral underlying stock is modeled by*

$$\begin{aligned} \frac{dS_t}{S_t} &= rdt + \sqrt{v_t}dB_t^S, \\ dv_t &= \kappa(\theta - v_t)dt + \sigma_v\sqrt{v_t}dB_t^v, \end{aligned}$$

where two standard Brownian motions, B_t^S and B_t^v, are correlated with a constant coefficient ρ. The model-implied squared VIX and skewness at time t for time to maturity τ are as follows

$$VIX_{t,\tau}^2 = (1-\omega)\theta + \omega v_t, \qquad \omega = \frac{1-e^{-\kappa\tau}}{\kappa\tau}, \tag{12}$$

$$Sk_{t,\tau} = \frac{TC^H}{(Var^H)^{3/2}}, \tag{13}$$

where the term variance and third cumulant are given by

$$Var^H = E_t(X_T^2) - E_t(X_TY_T) + \frac{1}{4}E_t(Y_T^2), \tag{14}$$

$$TC^H = E_t(X_T^3) - \frac{3}{2}E_t(X_T^2Y_T) + \frac{3}{4}E_t(X_TY_T^2) - \frac{1}{8}E_t(Y_T^3). \tag{15}$$

The integrated return uncertainty, X_T, and integrated instantaneous variance uncertainty, Y_T, are defined by

$$X_T \equiv \int_t^T \sqrt{v_u}dB_u^S, \qquad Y_T \equiv \int_t^T [v_u - E_t(v_u)]du = \sigma_v \int_t^T \frac{1-e^{-\kappa(T-u)}}{\kappa}\sqrt{v_u}dB_u^v.$$

The variance and covariance of X_T and Y_T are given by

$$E_t(X_T^2) = \int_t^T E_t(v_u)du, \tag{16}$$

$$E_t(X_TY_T) = \rho\sigma_v \int_t^T A_1E_t(v_u)du, \qquad A_1 = \frac{1-e^{-\kappa\tau^*}}{\kappa}, \qquad \tau^* = T-u, \tag{17}$$

$$E_t(Y_T^2) = \sigma_v^2 \int_t^T A_1^2E_t(v_u)du. \tag{18}$$

The third and co-third cumulants of X_T and Y_T are given by

$$E_t(X_T^3) = 3\rho\sigma_v \int_t^T A_1 \ E_t(v_u)du, \tag{19}$$

$$\begin{aligned} E_t(X_T^2 Y_T) &= \sigma_v^2 \int_t^T A_2 \ E_t(v_u)du, \\ A_2 &= \left(\frac{1-e^{-\kappa\tau^*}}{\kappa}\right)^2 + 2\rho^2 \frac{1-e^{-\kappa\tau^*}-\kappa\tau^* e^{-\kappa\tau^*}}{\kappa^2}, \end{aligned} \tag{20}$$

$$\begin{aligned} E_t(X_T Y_T^2) &= \rho\sigma_v^3 \int_t^T A_3 \ E_t(v_u)du, \\ A_3 &= 2\frac{1-e^{-\kappa\tau^*}-\kappa\tau^* e^{-\kappa\tau^*}}{\kappa^2}\frac{1-e^{-\kappa\tau^*}}{\kappa} + \frac{1-e^{-2\kappa\tau^*}-2\kappa\tau^* e^{-\kappa\tau^*}}{\kappa^3}, \end{aligned} \tag{21}$$

$$\begin{aligned} E_t(Y_T^3) &= 3\sigma_v^4 \int_t^T A_4 \ E_t(v_u)du, \\ A_4 &= \frac{1-e^{-2\kappa\tau^*}-2\kappa\tau^* e^{-\kappa\tau^*}}{\kappa^3}\frac{1-e^{-\kappa\tau^*}}{\kappa}, \end{aligned} \tag{22}$$

and $E_t(v_u) = \theta + (v_t - \theta)e^{-\kappa(u-t)}$ is the expected instantaneous variance.

Proof. See Zhang et al (2015).

Remark 3.1. Through asymptotic analysis, Zhang et al (2015) shows that for $\kappa > 0$, if τ is small, the skewness is given by

$$Sk_{t,\tau} = \frac{3}{2}\rho\frac{\sigma_v}{\sqrt{v_t}}\sqrt{\tau} + o(\sqrt{\tau}).$$

If τ is large, then the skewness is given by

$$Sk_{t,\tau} = \frac{3\rho\frac{\sigma_v}{\kappa} - \frac{3}{2}\frac{\sigma_v^2}{\kappa^2}}{\sqrt{1-\rho\frac{\sigma_v}{\kappa}+\frac{1}{4}\frac{\sigma_v^2}{\kappa^2}}}\frac{1}{\sqrt{\theta\tau}} + o\left(\frac{1}{\sqrt{\tau}}\right).$$

Hence the regularized skewness $\sqrt{\tau}\ Sk_{t,\tau}$ behaves linearly with τ for small τ, approaches to a constant for large τ.

When modeling the VIX term structure, Luo and Zhang (2012) observe that both short and long ends of the term structure are time-varying. Hence it is necessary to include a new factor of stochastic long-term mean into the standard Heston model in order to enhance its performance in pricing options. The impact of the new factor on skewness is presented in the next proposition.

Proposition 4 *In a modified Heston model with stochastic long term mean, the risk-neutral underlying stock is modeled by*

$$\begin{aligned} \frac{dS_t}{S_t} &= rdt + \sqrt{v_t}dB_t^S \\ dv_t &= \kappa(\theta_t - v_t)dt + \sigma_v\sqrt{v_t}dB_t^v. \\ d\theta_t &= \sigma_\theta\sqrt{\theta_t}dB_t^\theta, \end{aligned}$$

where the new standard Brownian motion, B_t^θ, is independent of B_t^S and B_t^v. The model-implied squared VIX and skewness at time t for time to maturity τ are as follows

$$\begin{aligned} VIX_{t,\tau}^2 &= (1-\omega)\theta_t + \omega v_t, \qquad \omega = \frac{1-e^{-\kappa\tau}}{\kappa\tau}, \\ Sk_{t,\tau} &= \frac{TC^H + TC^{HM}}{[Var^H + Var^{HM}]^{3/2}}, \end{aligned} \tag{23}$$

where the contributions of long term mean variation to the variance and the third cumulant are given by

$$Var^{HM} = \frac{1}{4}\sigma_\theta^2\theta_t C_1, \tag{24}$$

$$TC^{HM} = -\frac{3}{2}\sigma_\theta^2\theta_t C_1 + \frac{3}{2}\rho\sigma_v\sigma_\theta^2\theta_t C_2 - \frac{3}{8}\sigma_v^2\sigma_\theta^2\theta_t C_3 - \frac{3}{8}\sigma_\theta^4\theta_t C_4, \tag{25}$$

$$C_1 = \int_t^T \frac{(e^{-\kappa(T-u)} - 1 + \kappa(T-u))^2}{\kappa^2}du, \tag{26}$$

$$C_2 = \int_t^T \frac{1-e^{-\kappa(T-u)}}{\kappa}\int_t^u (1-e^{-\kappa(u-s)})\frac{e^{-\kappa(T-s)} - 1 + \kappa(T-s)}{\kappa}dsdu, \tag{27}$$

$$C_3 = \int_t^T \frac{(1-e^{-\kappa(T-u)})^2}{\kappa^2}\int_t^u (1-e^{-\kappa(u-s)})\frac{e^{-\kappa(T-s)} - 1 + \kappa(T-s)}{\kappa}dsdu, \tag{28}$$

$$C_4 = \int_t^T \frac{(e^{-\kappa(T-u)} - 1 + \kappa(T-u))^2}{\kappa^2}\int_t^u \frac{e^{-\kappa(T-s)} - 1 + \kappa(T-s)}{\kappa}dsdu, \tag{29}$$

and the term variance, Var^H, and third cumulant, TC^H, of Heston model are given by Proposition 3.

Proof. See Appendix B.2.

Remark 4.1. In modified Heston model, for $\kappa > 0$, if τ is small, the asymptotic skewness is the same as that in Heston model, given by

$$Sk_{t,\tau} = \frac{3}{2}\rho\frac{\sigma_v}{\sqrt{v_t}}\sqrt{\tau} + o(\sqrt{\tau}).$$

If τ is large, then the skewness is given by

$$Sk_{t,\tau} = -\frac{3\sqrt{3}}{5}\frac{\sigma_\theta}{\sqrt{\theta_t}}\sqrt{\tau} + o(\sqrt{\tau}).$$

Hence the regularized skewness $\sqrt{\tau}\ Sk_{t,\tau}$ behaves linearly with τ for both small τ and large τ.

With Merton (1976) jump-diffusion model, we are not able to create a flexible SKEW term structure because the model implied skewness times $\sqrt{\tau}$ is a constant across different τ. With Heston (1993) model, we are not able to produce a large short-term skewness because the model implied skewness goes to zero for small τ. Hence it is necessary to combine these two models in order to create a SKEW term structure flexible enough to fit market data. The result of a hybrid Merton-Heston model is presented in the next proposition.

Proposition 5 *In a hybrid Merton-Heston model, i.e., jump-diffusion model with stochastic volatility, the risk-neutral underlying stock is modeled by*

$$\begin{aligned}
\frac{dS_t}{S_t} &= rdt + \sqrt{v_t}dB_t^S + (e^x - 1)dN_t - \lambda(e^x - 1)dt, \\
dv_t &= \kappa(\theta - v_t)dt + \sigma_v\sqrt{v_t}dB_t^v.
\end{aligned}$$

The model-implied squared VIX and skewness at time t for time to maturity τ are as follows

$$\begin{aligned} VIX^2_{t,\tau} &= (1-\omega)\theta + \omega v_t + 2\lambda(e^x - 1 - x), \qquad \omega = \frac{1-e^{-\kappa\tau}}{\kappa\tau}, \\ Sk_{t,\tau} &= \frac{TC^H + \lambda x^3 \tau}{[Var^H + \lambda x^2 \tau]^{3/2}}, \end{aligned} \tag{30}$$

where the term variance, Var^H, and third cumulant, TC^H, of Heston model are given by Proposition 3.

Remark 5.1. Due to the independence between jump and diffusion processes, the term variance and third-cumulant of the hybrid Merton-Heston model is simply the sum of the contributions from each model. The short-term skewness is no longer zero due to the third cumulant contributed from jumps.

Remark 5.2. In hybrid Merton-Heston model, for $\kappa > 0$, if τ is small, the asymptotic skewness is given by

$$Sk_{t,\tau} = \frac{\lambda x^3}{(v_t + \lambda x^2)^{3/2}} \frac{1}{\sqrt{\tau}} + c\sqrt{\tau} + o(\sqrt{\tau}).$$

where

$$c = \frac{3\rho\sigma_v v_t}{2(v_t + \lambda x^2)^{3/2}} + \frac{3[\rho\sigma_v v_t + \kappa(v_t - \theta^v)]\lambda x^3}{4(v_t + \lambda x^2)^{5/2}}.$$

If τ is large, then the skewness is given by

$$Sk_{t,\tau} = \frac{a}{b\sqrt{b}} \frac{1}{\sqrt{\tau}} + o\left(\frac{1}{\sqrt{\tau}}\right),$$

where

$$\begin{aligned} a &= \lambda x^3 + \theta \left[3\rho \frac{\sigma_v}{\kappa} - \frac{3}{2}(1+2\rho^2)\frac{\sigma_v^2}{\kappa^2} + \frac{9}{4}\rho\frac{\sigma_v^3}{\kappa^3} - \frac{3}{8}\frac{\sigma_v^4}{\kappa^4} \right], \\ b &= \lambda x^2 + \theta\left(1 - \rho\frac{\sigma_v}{\kappa} + \frac{1}{4}\frac{\sigma_v^2}{\kappa^2}\right). \end{aligned}$$

Hence the regularized skewness $\sqrt{\tau}\ Sk_{t,\tau}$ approaches to a constant for both small τ and large τ.

The time-varying feature of VIX has been picked up by stochastic instantaneous variance, v_t, in the continuous-time models presented before. The time-varying feature of SKEW has to be picked up by jump-related variables. The jump size is usually assumed to follow a static distribution (in particular, a constant in this paper), hence we have to rely on stochastic jump intensity, λ_t, to capture the time-varying SKEW.

To begin with, we look at a simple case of mean-reverting deterministic intensity as follows.

Proposition 6 *In a jump-diffusion model with stochastic volatility and deterministic jump intensity with the same mean-reverting speed, the risk-neutral underlying stock is modeled by*

$$\begin{aligned} \frac{dS_t}{S_t} &= rdt + \sqrt{v_t}dB_t^S + (e^x - 1)dN_t - \lambda_t(e^x - 1)dt, \\ dv_t &= \kappa(\theta^v - v_t)dt + \sigma_v\sqrt{v_t}dB_t^v, \\ d\lambda_t &= \kappa(\theta^\lambda - \lambda_t)dt, \end{aligned}$$

The model-implied squared VIX and skewness at time t for time to maturity τ are as follows

$$\begin{aligned} VIX_{t,\tau}^2 &= (1-\omega)\theta^V + \omega V_t, \\ Sk_{t,\tau} &= \frac{TC^H + \Lambda_t^T x^3 \tau}{[Var^H + \Lambda_t^T x^2 \tau]^{3/2}}, \end{aligned} \tag{31}$$

where

$$\theta^V = \theta^v + 2\theta^\lambda(e^x - 1 - x), \qquad V_t = v_t + 2\lambda_t(e^x - 1 - x),$$

and the average jump intensity, Λ_t^T, is given by

$$\Lambda_t^T = \frac{1}{T-t}\int_t^T \lambda_u du = (1-\omega)\theta^\lambda + \omega\lambda_t.$$

The term variance, Var^H, and third cumulant, TC^H, of Heston model are given by Proposition 3.

Remark 6.1. The time-varying deterministic jump intensity does not create any difficulty in computing the term skewness. To produce a corrected result, we can simply replace constant jump intensity in the formula of Proposition 5 by the average one over the term period.

Remark 6.2. The mean-reverting speed of λ_t is designed to be the same as that of v_t, so that the resulted *instantaneous squared VIX* defined by Luo and Zhang (2012), $V_t = v_t + 2\lambda_t(e^x - 1 - x)$, follows a mean-reverting process with the same speed

$$dV_t = \kappa(\theta^V - V_t)dt + \sigma_v\sqrt{v_t}dB_t^v.$$

The VIX term structure model of Luo and Zhang (2012) can be directly applied.

With these background knowledge, we are now ready to move on to a more complicated model of stochastic jump intensity as follows.

Proposition 7 *In a jump-diffusion model with stochastic volatility and stochastic jump intensity with the same mean-reverting speed, the risk-neutral underlying stock is modeled by*

$$\begin{aligned}
\frac{dS_t}{S_t} &= rdt + \sqrt{v_t}dB_t^S + (e^x - 1)dN_t - \lambda_t(e^x - 1)dt, \\
dv_t &= \kappa(\theta^v - v_t)dt + \sigma_v\sqrt{v_t}dB_t^v, \\
d\lambda_t &= \kappa(\theta^\lambda - \lambda_t)dt + \sigma_\lambda\sqrt{\lambda_t}dB_t^\lambda,
\end{aligned}$$

where B_t^λ is independent of B_t^S, B_t^v, N_t. The model-implied squared VIX and skewness *at time t for time to maturity τ are as follows*

$$
\begin{aligned}
VIX_{t,\tau}^2 &= (1-\omega)\theta^V + \omega V_t, \\
Sk_{t,\tau} &= \frac{TC^H + TC^J}{[Var^H + Var^J]^{3/2}},
\end{aligned} \tag{32}
$$

where the contributions of jump component to variance and third cumulant are given by

$$
Var^J = \Lambda_t^T x^2 \tau + (e^x - 1 - x)^2 E_t(Z_T^2), \tag{33}
$$

$$
TC^J = \Lambda_t^T x^3 \tau - 3x^2(e^x - 1 - x)E_t(Z_T^2) - (e^x - 1 - x)^3 E_t(Z_T^3), \tag{34}
$$

Z_T is defined by

$$
Z_T \equiv \int_t^T [\lambda_u - E_t(\lambda_u)]du = \sigma_\lambda \int_t^T \frac{1 - e^{-\kappa(T-u)}}{\kappa} \sqrt{\lambda_u} dB_u^\lambda,
$$

and the variance and third cumulant of Z_T are given by

$$
E_t(Z_T^2) = \sigma_\lambda^2 \int_t^T \frac{(1 - e^{-\kappa\tau^*})^2}{\kappa^2} E_t(\lambda_u)du, \qquad \tau^* = T - u, \tag{35}
$$

$$
E_t(Z_T^3) = 3\sigma_\lambda^4 \int_t^T \frac{1 - e^{-2\kappa\tau^*} - 2\kappa\tau^* e^{-\kappa\tau^*}}{\kappa^3} \frac{1 - e^{-\kappa\tau^*}}{\kappa} E_t(\lambda_u)du, \tag{36}
$$

and $E_t(\lambda_u) = \theta^\lambda + (v_t - \theta^\lambda)e^{-\kappa(u-t)}$ is the expected jump intensity. The average jump intensity Λ_t^T is given by Proposition 6, and the term variance, Var^H, and third cumulant, TC^H, of Heston model are given by Proposition 3.

Proof. See Appendix B.3.

Remark 7.1. As we can see from equations (33) and (34), the term variance, Var^J, and third cumulant, TC^J, contributed from jumps consist of two components. One of them is due to the average jump intensity, Λ_t^T; the other one is due to the uncertainty, σ_λ, in jump intensity. There is no interaction term between stochastic volatility and jumps because they are independent.

Remark 7.2. In this three-factor model, for $\kappa > 0$, if τ is small, the asymptotic skewness is given by

$$Sk_{t,\tau} = \frac{\lambda_t x^3}{(v_t + \lambda_t x^2)^{3/2}} \frac{1}{\sqrt{\tau}} + c\sqrt{\tau} + o(\sqrt{\tau}),$$

where

$$c = \frac{3\rho\sigma_v v_t}{2(v_t + \lambda_t x^2)^{3/2}} + \frac{(3\rho\sigma_v v_t \lambda_t + \kappa v_t \lambda_t + 2\kappa v_t \theta^\lambda - 3\kappa\theta^v \lambda_t)x^3}{4(v_t + \lambda_t x^2)^{5/2}} + \frac{\kappa\lambda_t(\lambda_t - \theta^\lambda)x^5}{4(v_t + \lambda_t x^2)^{5/2}}.$$

If τ is large, then the skewness is given by

$$Sk_{t,\tau} = \frac{a}{b\sqrt{b}} \frac{1}{\sqrt{\tau}} + o\left(\frac{1}{\sqrt{\tau}}\right),$$

where

$$\begin{aligned} a &= \theta^\lambda \left[x^3 - 3x^2(e^x - 1 - x)\frac{\sigma_\lambda^2}{\kappa^2} - 3(e^x - 1 - x)^3 \frac{\sigma_\lambda^4}{\kappa^4} \right] \\ &\quad + \theta^v \left[3\rho\frac{\sigma_v}{\kappa} - \frac{3}{2}(1 + 2\rho^2)\frac{\sigma_v^2}{\kappa^2} + \frac{9}{4}\rho\frac{\sigma_v^3}{\kappa^3} - \frac{3}{8}\frac{\sigma_v^4}{\kappa^4} \right], \\ b &= \theta^\lambda \left[x^2 + (e^x - 1 - x)^2 \frac{\sigma_\lambda^2}{\kappa^2} \right] + \theta^v \left(1 - \rho\frac{\sigma_v}{\kappa} + \frac{1}{4}\frac{\sigma_v^2}{\kappa^2} \right). \end{aligned}$$

Hence the regularized skewness $\sqrt{\tau}\ Sk_{t,\tau}$ approaches to a constant for both small τ and large τ.

As explained early, we need two variance factors, i.e., instantaneous one, v_t, and long-term one, θ_t^v to capture the information of time-varying VIX term structure. Similarly, in order the capture the information of time-varying SKEW term structure, we also need two factors, i.e., instantaneous and long-term jump intensity, λ_t and θ_t^λ.

The simplest five-factor model is presented as follows.

Proposition 8 *In a jump-diffusion model with stochastic volatility and stochastic jump intensity as well as stochastic corresponding long term mean levels, the risk-*

neutral underlying stock is modeled by

$$\begin{aligned} \frac{dS_t}{S_t} &= rdt + \sqrt{v_t}dB_t^S + (e^x - 1)dN_t - \lambda_t(e^x - 1)dt, \\ dv_t &= \kappa(\theta_t^v - v_t)dt + \sigma_v\sqrt{v_t}dB_t^v, \\ d\lambda_t &= \kappa(\theta_t^\lambda - \lambda_t)dt + \sigma_\lambda\sqrt{\lambda_t}dB_t^\lambda, \\ d\theta_t^v &= \sigma_1\sqrt{\theta_t^v}dB_t^{\theta,v}, \\ d\theta_t^\lambda &= \sigma_2\sqrt{\theta_t^\lambda}dB_t^{\theta,\lambda}, \end{aligned}$$

where $B_t^{\theta,v}$, $B_t^{\theta,\lambda}$ *are independent of each other and* B_t^S, B_t^v, B_t^λ *and* N_t. *The model-implied squared VIX and skewness at time* t *for time to maturity* τ *are as follows*

$$\begin{aligned} VIX_{t,\tau}^2 &= (1-\omega)\theta^V + \omega V_t, \\ Sk_{t,\tau} &= \frac{TC^H + TC^J + TC^M}{[Var^H + Var^J + Var^M]^{3/2}}. \end{aligned} \tag{37}$$

where the contributions of long term mean variation to variance and third cumulant are given by

$$Var^M = Var^{HM} + (e^x - 1 - x)^2\sigma_2^2\theta_t^\lambda C_1, \tag{38}$$

$$TC^M = TC^{HM} - 3\sigma_2^2\theta_t^\lambda[x^2(e^x - 1 - x)C_1 + (e^x - 1 - x)^3(\sigma_\lambda^2 C_3 + \sigma_2^2 C_4)], \tag{39}$$

and the variance Var^H, *third cumulant* TC^H, *of Heston model are given by Proposition 3,* Var^{HM}, TC^{HM}, C_1, C_3, C_4 *are given by Proposition 4,* Var^J *and* TC^J, *of jump component are given by Proposition 7.*

Remark 8.1. The result is built by combining those of Propositions 4 and 7 and including additional term variance and third cumulant contributed from the uncertainty of long-term mean jump intensity, θ_t^λ.

Remark 8.2. In this five-factor model, for $\kappa > 0$, if τ is small, the asymptotic skewness is given by

$$Sk_{t,\tau} = \frac{\lambda_t x^3}{(v_t + \lambda_t x^2)^{3/2}} \frac{1}{\sqrt{\tau}} + o\left(\frac{1}{\sqrt{\tau}}\right).$$

If τ is large, then the skewness is given by

$$Sk_{t,\tau} = \frac{a}{b\sqrt{b}}\sqrt{\tau} + o(\sqrt{\tau}),$$

where

$$\begin{aligned} a &= -\frac{1}{5}(e^x - 1 - x)^3 \theta_t^\lambda \sigma_2^4 - \frac{1}{40}\theta_t^v \sigma_1^4, \\ b &= \frac{1}{3}(e^x - 1 - x)^2 \theta_t^\lambda \sigma_2^2 + \frac{1}{12}\theta_t^v \sigma_1^2. \end{aligned}$$

Hence the regularized skewness $\sqrt{\tau}\ Sk_{t,\tau}$ approaches to a constant for small τ, and behaves linearly with τ for large τ.

4 Data

The daily data on two term structures of VIX and SKEW ranging from 2 January 1990 to 31 December 2014 are provided by the CBOE.

On each day, we have VIX and SKEW data for up to ten maturity dates. The time to maturity becomes one day shorter as we move forward by one day. In order to facilitate our estimation of continuous-time models, we construct a new set of data with constant times to maturity from the original one.

We are particularly interested in times to maturity, τ = 1, 3, 6, 9, 12 and 15 months. Following the practice of the CBOE, we compute the SKEW at time t for time to maturity τ by using interpolation as follows

$$SKEW_{t,\tau} = \omega SKEW_{t,\text{last}} + (1-\omega) SKEW_{t,\text{next}},$$

where ω is a weight determined by

$$\omega = \frac{\tau_{\text{next}} - \tau}{\tau_{\text{next}} - \tau_{\text{last}}},$$

and τ_{last} and τ_{next} are the times to maturity[2] (up to minute) of the last and next available data respectively. We set SKEW term structure to be flat if τ is larger than the longest available time to maturity.

Figure 1 shows the time evolution of SKEW term structure for time to maturity from 1 month to 15 months and Figure 2 shows the time series of three selected SKEWs with constant time to maturity. The sample period is from 2 January 1990 to 31 December 2014. As we can see from the figures, the SKEW term structure has a relative stable shape before financial tsunami. It becomes erratic in the last five years. The level of SKEW significantly increases in recent years. It indicates that option traders expect higher tail risk after the financial crisis.

Figure 3 shows a few samples of SKEW term structure with abnormal shape. They either have a very low minimum SKEW (smaller than 90) or a very high maximum SKEW (higher than 180).

5 Model Estimation and Its Empirical Performance

One always encounters the challenge that the daily realizations of diffusive volatility or jump intensity are unobservable. To circumvent the estimation difficulty of latent variables, we use a two-stage iterative approach. In the first stage, we follow the same procedure adopted by Luo and Zhang (2012) to estimate the mean-reverting speed κ and daily realizations of instantaneous squared VIX (ISVIX) and its long term mean

[2]There is an inconsistency in the CBOE practice on calculating number of days for the time to maturity. Business day convention is used in computing the term structure of the SKEW, however calendar day is used in computing 30-day SKEW index. In this paper, we construct constant time to maturity SKEW by using business day convention.

level using the term structure of CBOE VIX. Besides the VIX term structure data from 2 January 1992 to 4 September 2009, we extend the sample period including the latest VIX term structure data from 24 November 2010 to 31 December 2014. In the second stage, we first calibrate the correlation coefficient ρ and jump size x, and then estimate the volatility of volatility σ_v as well as the daily realizations of Brownian variance and its long term mean level using the term structure of CBOE SKEW. Meanwhile, we obtain the daily realizations of jump intensity and its long term mean level due to the linear relationship of instantaneous squared VIX to Brownian variance and jump intensity.

For estimation accuracy, we select the matched term structure data of VIX and SKEW both with no less than 6 time to maturities. We design a 5-step estimation procedure as follows:

Step 1 Given an initial value κ, obtain the time series of instantaneous and long-term value of ISVIX $\{V_t, \theta_t^V\}$, $t = 1, 2, \cdots, T$, where T is the total number of trading days in the matched sample. In this step, we are required to solve T optimization problems

$$\{V_t, \theta_t^V\} = argmin \sum_{j=1}^{n_t} (VIX_{t,\tau_j} - VIX_{t,\tau_j}^{Mkt})^2,$$

where VIX_{t,τ_j} is the model-implied value of VIX with maturity τ_j on day t, VIX_{t,τ_j}^{Mkt} is the corresponding market value, n_t is the number of maturities on day t.

Step 2 Estimate κ with $\{V_t, \theta_t^V\}$ obtained in step 1 by minimizing the overall objective function for VIX:

$$\{\kappa\} = argmin \sum_{t=1}^{T} \sum_{j=1}^{n_t} (VIX_{t,\tau_j} - VIX_{t,\tau_j}^{Mkt})^2.$$

Step 1 and Step 2 are repeated to estimate $\{\kappa\}$ until no further significant decrease exists in the overall objective (i.e., the aggregate sum of squared VIX errors).

Step 3 Determine the correlation coefficient ρ and the jump size x externally using the following relations:

$$Corr_t(d\log S_t, dVIX_t^2) \approx \rho\sqrt{\frac{v_t}{v_t + x^2\lambda_t}},$$
$$RSk_{t,\tau\to 0} = \frac{\lambda_t x^3}{(v_t + \lambda_t x^2)^{3/2}},$$

where $Corr_t(d\log S_t, dVIX_t^2)$ denotes the conditional correlation between the logarithmic return and the daily change of VIX square neglecting the instantaneous variance of λ_t, θ_t^v and θ_t^λ, and $RSk_{t,\tau\to 0}$ denotes the daily instantaneous regularized skewness.

Step 4 With the optimal mean-reverting speed κ and time series $\{V_t, \theta_t^V\}$ obtained in the first two steps as well as the correlation coefficient ρ and jump size x in Step 3, given an initial value σ_v, we can estimate daily Brownian variance v_t by solving T optimizations

$$\{v_t, \theta_t^v\} = argmin\sum_{j=1}^{n_t}(SKEW_{t,\tau_j} - SKEW_{t,\tau_j}^{Mkt})^2,$$

with constraints $v_t < V_t$ and $\theta_t^v < \theta_t^V$, where $SKEW_{t,\tau_j}$ is the theoretical value of SKEW with maturity τ_j on date t, $SKEW_{t,\tau_j}^{Mkt}$ is the corresponding market value.

Step 5 Estimate σ_v with v_t obtained in step 4 by minimizing the overall objective function for SKEW:

$$\{\sigma_v\} = argmin\sum_{t=1}^{T}\sum_{j=1}^{n_t}(SKEW_{t,\tau_j} - SKEW_{t,\tau_j}^{Mkt})^2.$$

Step 4 and Step 5 are repeated to estimate σ_v until no further significant decrease exists in the overall objective (i.e., the aggregate sum of squared SKEW errors).

We conduct the estimation procedure using the 5-factor model from Proposition 8. The full sample optimal κ is 3.5680, as shown in Table 2, and the daily realizations of ISVIX and its long term mean level are shown in Figure 4. We also report the subperiod optimal values of κ as well as the mean values and standard deviations of ISVIX and its long term mean level. The mean reverting speed κ for 2 January 1992 to 4 September 2009 is 5.8285, whereas κ for 24 November 2010 to 31 December 2014 is 1.9304. During 2008 financial crisis, ISVIX changed dramatically, resulting in a large mean reverting speed. The post-2008 VIX term structure data exhibit higher long term values than the instantaneous.

We assume the aggregate proportion of Brownian motion variance to total variance is 0.7. The optimal average of V_t is 0.0438 from the full sample VIX term structure estimation. The correlation coefficient between VIX square and logarithmic SPX daily changes is -0.7185 (sample period 2 January 1990 to 31 December 2014). We use the spline interpolation -0.2 of the first 4 weekly average regularized skewness -0.2817, -0.3786, -0.4759, -0.5430 with 7, 12, 17, 22 business days to maturity respectively (sample period 2 January 1990 to 31 December 2014) as the limit of regularized skewness at point zero. Given the above conditions, we have jump size x and correlation coefficient ρ as -0.14, -0.859 respectively.

Given the mean reverting speed κ, correlation coefficient ρ, jump size x as well as the daily realizations of ISVIX and its long term mean level, the optimal value of the volatility of volatility σ_v is 0.7929 using the SKEW formula implied in Model 8 ignoring the instantaneous variance of jump intensity λ_t, Brownian variance long term mean θ_t^v and jump intensity long term mean θ_t^λ. The daily realizations of Brownian variance and jump intensity as well as their long term mean levels are shown in Figure 5 and Figure 6 respectively. We see that the increase of Brownian variance tends to be associated with the increase of jump intensity during financial crises. The relative jump risk indicators, the proportion of Brownian variance to total variance and the instantaneous regularized skewness, are shown in Figure 7. The aggregate estimation performance is shown in Figure 9 and Figure 10.

Note that we omit the instantaneous variance of λ_t, θ_t^v and θ_t^λ. After getting the time series of Brownian variance v_t, jump intensity λ_t as well as their long term mean levels θ^v and θ_t^λ, it is straightforward to calculate the instantaneous volatility of λ_t, θ_t^λ and θ_t^v as 11.1370, 3.5402 and 0.5362 respectively through the variance of daily increments. As the expected $\mathcal{F}_t$-conditional variance $\mathbb{E}(var_t(d\lambda_t)) = var(d\lambda_t) - var(\mathbb{E}_t(d\lambda_t))$, the second term on the right-hand side is negligible due to its order $(dt)^2$, where $dt = \frac{1}{252}$ neglecting the single point effect of data unavailability during 2009-2010. The expected conditional variance is exactly the unconditional variance for the long term mean levels due to the martingale property.

The correlation coefficient between the estimated Brownian variance and jump intensity is 0.055, see Figure 8, which indicates the jump intensity is neither constant nor proportional to Brownian variance as assumed in Bates (1996) and Bates (2000). However, the models in this paper retain the affine structure in Duffie, Pan and Singleton (2000). Thus, the daily realization of Brownian variance and jump intensity can be directly applied to option pricing.

We show in the model section that Black-Scholes (1973), Merton (1976), Heston (1993) models exhibit zero skewness, constant regularized skewness and zero instantaneous skewness, respectively, which are implausible to explain the stylized skewness in the distribution of the SPX log returns, see Das and Sundaram (1999) and Zhang et al. (2015). Note that the nonzero conditional correlation between the SPX logrithmic return and the daily change of VIX is dominated by the Brownian variance term. It is essential to incorporate the Heston stochastic volatility into the model setup. Luo and Zhang (2012) show advantages of modelling the term structure of VIX using an extended Heston stochastic volatility model, where the long term mean level follows a martingale process. Therefore, within the framework of Luo and Zhang (2012), we compare the empirical performance of the extended stochastic volatility models with constant, deterministic and stochastic jump intensities, abbreviated as SVCJ, SVDJ, SVSJ, respectively. Note that the constant jump variance, which is the jump intensity multiplied by a quadratic function of the constant jump size in our model setting, is less than the total instantaneous squared VIX over the entire sample period. The minimum value, 32.7 basis points, of the estimated optimal time series of instantaneous squared VIX indicates that the constant jump intensity is less than 0.02, which is too small to be sensible in the SVCJ model as a jump arrives per 50 years. Similar argument holds for the SVDJ model as the minimum value of long term squared VIX is 3.9 basis points. Thus, the fitting performance supports the 5-factor model SVSJ in Proposition 8 over the SVCJ and SVDJ models.

6 Conclusion

In this paper, we derive skewness formulas under various affine jump diffusion models, proposed by Duffie, Pan and Singleton (2000), which are typically used in option pricing. Given the VIX formulas in Luo and Zhang (2012), the skewness formulas provide a new perspective to identify the model parameters as well as latent variables using the CBOE VIX and SKEW term structure data, which are daily updated on the CBOE website, as opposed to the option cross-sectional data, which are only available in subscribed databases. We also analyse the asymptotic behaviors of the skewness formulas.

To model skewness more accurately, we propose an affine jump diffusion model with 5 state variables, which are log returns, Brownian variance, jump intensity and the long term mean levels of Brownian variance and jump intensity. We adopt a two-stage iterative procedure to estimate the parameters as well as the latent Brownian variance and jump intensity processes. As the VIX and SKEW term structure data are extracted from option prices, the parameters and latent variables are estimated under risk-neutral measure, and can be directly applied to option pricing. The Brownian variance and jump variance can also be used to explain stock returns, and quantify the variance risk premium and jump risk premium. The CBOE SKEW, as a complementary index for VIX, is informative for both model specification and risk quantification.

A The First Three Moments of Log Return

Bakshi, Kapadia and Madan (2003) propose a methodology of evaluating the first three moments of log return, $E_t^Q(R_t^T)$, $E_t^Q[(R_t^T)^2]$ and $E_t^Q[(R_t^T)^3]$ by using current prices of European options as follows.

For any twice differentiable function $f(S_T)$, following equality holds

$$
\begin{aligned}
f(S_T) &= f(K_0) + f'(K_0)(S_T - K_0) + \int_0^{K_0} f''(K) \max(K - S_T, 0) dK \\
&\quad + \int_{K_0}^{+\infty} f''(K) \max(S_T - K, 0) dK,
\end{aligned} \tag{40}
$$

where K_0 is a reference strike price that could take any value. This mathematical equality has a profound financial meaning: A European-style derivative with an arbitrary payoff function, $f(S_T)$, can be decomposed into a portfolio of bond with face value $f(K_0)$, $f'(K_0)$ amount of forward contract and $f''(K)$ amount of European options with strikes between 0 and K_0 for puts and between K_0 and $+\infty$ for calls.

Applying equation (40) to the power of log return gives

$$
\begin{aligned}
\ln \frac{S_T}{F_t^T} &= \ln \frac{K_0}{F_t^T} + \frac{S_T}{K_0} - 1 - \int_0^{K_0} \frac{1}{K^2} \max(K - S_T, 0) dK \\
&\quad - \int_{K_0}^{+\infty} \frac{1}{K^2} \max(S_T - K, 0) dK, \\
\ln^2 \frac{S_T}{F_t^T} &= \ln^2 \frac{K_0}{F_t^T} + 2 \ln \frac{K_0}{F_t^T} \left(\frac{S_T}{K_0} - 1 \right) + \int_0^{K_0} \frac{2 - 2 \ln \frac{K}{F_t^T}}{K^2} \max(K - S_T, 0) dK \\
&\quad + \int_{K_0}^{+\infty} \frac{2 - 2 \ln \frac{K}{F_t^T}}{K^2} \max(S_T - K, 0) dK, \\
\ln^3 \frac{S_T}{F_t^T} &= \ln^3 \frac{K_0}{F_t^T} + 3 \ln^2 \frac{K_0}{F_t^T} \left(\frac{S_T}{K_0} - 1 \right) + \int_0^{K_0} \frac{6 \ln \frac{K}{F_t^T} - 3 \ln^2 \frac{K}{F_t^T}}{K^2} \max(K - S_T, 0) dK \\
&\quad + \int_{K_0}^{+\infty} \frac{6 \ln \frac{K}{F_t^T} - 3 \ln^2 \frac{K}{F_t^T}}{K^2} \max(S_T - K, 0) dK.
\end{aligned}
$$

Applying conditional expectation to the three equations in risk-neutral measure, we notice that

$$E_t^Q[\max(S_T - K, 0)] = e^{r\tau} c_t(K), \qquad E_t^Q[\max(K - S_T, 0)] = e^{r\tau} p_t(K),$$

where $c_t(K)/p_t(K)$ is call/put option price at current time. Evaluating the integration approximately using discretization gives the first three moments in equations (5), (6) and (7).

B Model-implied VIX and Skewness Formulas

B.1 Merton Model

In Merton (1976) model, the risk-neutral logarithmic process of the underlying stock is as follows

$$d\ln S_t = \left[r - \frac{1}{2}\sigma^2 - \lambda(e^x - 1 - x)\right]dt + \sigma dB_t + x dN_t - \lambda x dt.$$

We define the logarithmic return from time t to T as $R_t^T \equiv \ln \frac{S_T}{S_t}$, then the variance and third cumulant of R_t^T are given by

$$\begin{aligned} E_t^Q[R_t^T - E_t^Q(R_t^T)]^2 &= E_t^Q\Big(\int_t^T \sigma dB_t + x dN_t - \lambda x dt\Big)^2 = (\sigma^2 + \lambda x^2)\tau, \\ E_t^Q[R_t^T - E_t^Q(R_t^T)]^3 &= E_t^Q\Big(\int_t^T \sigma dB_t + x dN_t - \lambda x dt\Big)^3 = \lambda x^3 \tau. \end{aligned}$$

Therefore, the skewness of logarithmic return is given by

$$Sk_t = \frac{E_t^Q[R_t^T - E_t^Q(R_t^T)]^3}{\{E_t^Q[R_t^T - E_t^Q(R_t^T)]^2\}^{3/2}} = \frac{\lambda x^3}{\sigma^2 + \lambda x^2}\tau^{-1/2}.$$

B.2 Heston Model with Stochastic Long Term Mean

We generalize the Heston variance process by adding another stochastic component, the long term mean θ_t, as follows

$$dv_t = \kappa(\theta_t - v_t)dt + \sigma_v \sqrt{v_t} dB_t^v, \tag{41}$$

$$d\theta_t = \sigma_\theta \sqrt{\theta_t} dB_t^\theta, \tag{42}$$

where B_t^θ is independent of B_t^v and B_t^S.

Converting equation (41) into stochastic integral form and plugging equation (42) yields

$$v_s = E_t(v_s) + \sigma_v \int_t^s e^{-\kappa(s-u)} \sqrt{v_u} dB_u^v + \sigma_\theta \int_t^s (1 - e^{-\kappa(s-u)}) \sqrt{\theta_u} dB_u^\theta, \qquad (43)$$

where the expectation of v_s at time t $(t < s)$ is given by

$$E_t(v_s) = \theta_t + (v_t - \theta_t) e^{-\kappa(s-t)}.$$

Following the procedure in Zhang et al (2015), we define two integrals as

$$X_T \equiv \int_t^T \sqrt{v_u} dB_u^S, \qquad Y_T \equiv \int_t^T [v_u - E_t^Q(v_u)] du.$$

Substituting equation (43) and interchanging the order of integrations gives

$$\begin{aligned} Y_T &= \sigma_v \int_t^T \int_t^s e^{-\kappa(s-u)} \sqrt{v_u} dB_u^v ds + \sigma_\theta \int_t^T \int_t^s (1 - e^{-\kappa(s-u)}) \sqrt{\theta_u} dB_u^\theta ds, \\ &= \sigma_v \int_t^T \frac{1 - e^{-\kappa(T-u)}}{\kappa} \sqrt{v_u} dB_u^v + \sigma_\theta \int_t^T \frac{e^{-\kappa(T-u)} - 1 + \kappa(T-u)}{\kappa} \sqrt{\theta_u} dB_u^\theta. \end{aligned}$$

We introduce new martingale processes, Y_s^H and Y_s^M, as follows

$$\begin{aligned} Y_s^H &= \sigma_v \int_t^s \frac{1 - e^{-\kappa(T-u)}}{\kappa} \sqrt{v_u} dB_u^v, \\ Y_s^M &= \sigma_\theta \int_t^s \frac{e^{-\kappa(T-u)} - 1 + \kappa(T-u)}{\kappa} \sqrt{\theta_u} dB_u^\theta. \end{aligned}$$

Therefore, at time T, we have $Y_T = Y_T^H + Y_T^M$.

To express the contribution of long term mean variation explicitly, using the independency of B_s^θ and the martingale property of Y_s^H, Y_s^M and θ_s, we expand and simplify the variance and third cumulant of R_t^T in Zhang et al (2015) as follows

$$\begin{aligned} E_t^Q[R_t^T - E_t^Q(R_t^T)]^2 &= E_t^Q(X_T^2) - E_t^Q(X_T Y_T) + \frac{1}{4} E_t^Q(Y_T^2) \\ &= Var^H - E_t^Q(X_T Y_T^M) + \frac{1}{4} E_t^Q[2 Y_T^H Y_T^M + (Y_T^M)^2] \\ &= Var^H + \frac{1}{4} E_t^Q[(Y_T^M)^2], \end{aligned}$$

$$
\begin{aligned}
E_t^Q[R_t^T - E_t^Q(R_t^T)]^3 &= E_t^Q(X_T^3) - \frac{3}{2}E_t^Q(X_T^2 Y_T) + \frac{3}{4}E_t^Q(X_T Y_T^2) - \frac{1}{8}E_t^Q(Y_T^3) \\
&= TC^H - \frac{3}{2}E_t^Q(X_T^2 Y_T^M) + \frac{3}{4}E_t^Q[2X_T Y_T^H Y_T^M + X_T(Y_T^M)^2] \\
&\quad -\frac{1}{8}E_t^Q[3(Y_T^H)^2 Y_T^M + 3Y_T^H(Y_T^M)^2 + (Y_T^M)^3] \\
&= TC^H - \frac{3}{2}E_t^Q(X_T^2 Y_T^M) + \frac{3}{2}E_t^Q(X_T Y_T^H Y_T^M) \\
&\quad -\frac{1}{8}E_t^Q[3(Y_T^H)^2 Y_T^M + (Y_T^M)^3],
\end{aligned}
$$

where the variance Var^H and third cumulant TC^H are of the original forms in Zhang et al (2015), with no impact of stochastic long term mean.

We need the following results in expressing the extra terms arising from the stochastic long term mean.

Lemma 1 *The correlations between Y_s^M and θ_s as well as v_s are given by*

$$
E_t^Q(Y_s^M \theta_s) = \sigma_\theta^2 \theta_t \int_t^s \frac{e^{-\kappa(T-u)} - 1 + \kappa(T-u)}{\kappa} du, \tag{44}
$$

$$
E_t^Q(Y_s^M v_s) = \sigma_\theta^2 \theta_t \int_t^s (1 - e^{-\kappa(s-u)}) \frac{e^{-\kappa(T-u)} - 1 + \kappa(T-u)}{\kappa} du. \tag{45}
$$

Proof. See Appendix B.2.1.

Using Ito's Isometry and the martingale property of θ_u, we have

$$
E_t^Q[(Y_T^M)^2] = E_t^Q\Big(\sigma_\theta \int_t^T \frac{e^{-\kappa(T-u)} - 1 + \kappa(T-u)}{\kappa} \sqrt{\theta_u} dB_u^\theta\Big)^2 = \sigma_\theta^2 \theta_t C_1,
$$

where $C_1 = \int_t^T \frac{(e^{-\kappa(T-u)} - 1 + \kappa(T-u))^2}{\kappa^2} du$.

Using Ito's Lemma and martingale property of X_u and Y_u^M, we have

$$
\begin{aligned}
E_t^Q(X_T^2 Y_T^M) &= E_t^Q \int_t^T d(X_u^2 Y_u^M) \\
&= E_t^Q \int_t^T 2X_u Y_u^M dX_u + X_u^2 dY_u^M + Y_u^M (dX_u)^2 + 2X_u dX_u dY_u^M \\
&= \int_t^T E_t^Q(Y_u^M v_u) du = \sigma_\theta^2 \theta_t C_1.
\end{aligned}
$$

Using Ito's Lemma and martingale property of X_u, Y_u^H and Y_u^M, we have

$$\begin{aligned} E_t^Q[X_T Y_T^H Y_T^M] &= E_t^Q \int_t^T d(X_u Y_u^H Y_u^M) \\ &= E_t^Q \int_t^T Y_u^H Y_u^M dX_u + X_u Y_u^M dY_u^H + X_u Y_u^H dY_u^M \\ &\quad + Y_u^M dX_u dY_u^H + Y_u^H dX_u dY_u^M + X_u dY_u^H dY_u^M \\ &= \rho\sigma_v \int_t^T \frac{1-e^{-\kappa(T-u)}}{\kappa} E_t^Q(Y_u^M v_u) du = \rho\sigma_v\sigma_\theta^2\theta_t C_2, \end{aligned}$$

where $C_2 = \int_t^T \frac{1-e^{-\kappa(T-u)}}{\kappa} \int_t^u (1-e^{-\kappa(u-s)}) \frac{e^{-\kappa(T-s)}-1+\kappa(T-s)}{\kappa} ds du.$

Using Ito's Lemma and martingale property of Y_u^H and Y_u^M, we have

$$\begin{aligned} E_t^Q[(Y_T^H)^2 Y_T^M] &= E_t^Q \int_t^T d[(Y_u^H)^2 Y_T^M] \\ &= E_t^Q \int_t^T 2Y_u^H Y_u^M dY_u^H + (Y_u^H)^2 dY_u^M + Y_u^M (dY_u^H)^2 + 2Y_u^H dY_u^H dY_u^M \\ &= \sigma_v^2 \int_t^T \frac{(1-e^{-\kappa(T-u)})^2}{\kappa^2} E_t^Q(Y_u^M v_u) du = \sigma_v^2 \sigma_\theta^2 \theta_t C_3, \end{aligned}$$

where $C_3 = \int_t^T \frac{(1-e^{-\kappa(T-u)})^2}{\kappa^2} \int_t^u (1-e^{-\kappa(u-s)}) \frac{e^{-\kappa(T-s)}-1+\kappa(T-s)}{\kappa} ds du.$

Using Ito's Lemma and martingale property of Y_u^M, we have

$$\begin{aligned} E_t^Q[(Y_T^M)^3] &= E_t^Q \int_t^T d(Y_u^M)^3 = E_t^Q \int_t^T 3(Y_u^M)^2 dY_u^M + 3Y_u^M (dY_u^M)^2 \\ &= 3\sigma_\theta^2 \int_t^T \frac{(e^{-\kappa(T-u)}-1+\kappa(T-u))^2}{\kappa^2} E_t^Q(Y_u^M \theta_u) du = 3\sigma_\theta^4 \theta_t C_4, \end{aligned}$$

where $C_4 = \int_t^T \frac{(e^{-\kappa(T-u)}-1+\kappa(T-u))^2}{\kappa^2} \int_t^u \frac{e^{-\kappa(T-s)}-1+\kappa(T-s)}{\kappa} ds du.$

B.2.1 Proof of Lemma 1

Using Ito's Lemma and martingale property of Y_u^M and θ_u, we have

$$\begin{aligned} E_t^Q(Y_s^M \theta_s) &= E_t^Q \int_t^s d(Y_u^M \theta_u) = E_t^Q \int_t^s \theta_u dY_u^M + Y_u^M d\theta_u + dY_u^M d\theta_u \\ &= \sigma_\theta^2 \theta_t \int_t^s \frac{e^{-\kappa(T-u)}-1+\kappa(T-u)}{\kappa} du, \end{aligned}$$

which is equivalent to equation (44).

Furthermore, we have the integral

$$\begin{aligned}\int_t^s E_t^Q(Y_u^M\theta_u)du &= \sigma_\theta^2\theta_t\int_t^s\int_t^u \frac{e^{-\kappa(T-r)}-1+\kappa(T-r)}{\kappa}drdu\\ &= \sigma_\theta^2\theta_t\int_t^s (s-u)\frac{e^{-\kappa(T-u)}-1+\kappa(T-u)}{\kappa}du.\end{aligned}$$

Using Ito's Lemma and martingale property of Y_u^M, we have

$$\begin{aligned}E_t^Q(Y_s^M v_s) &= E_t^Q\int_t^s d(Y_u^M v_u) = E_t^Q\int_t^s v_u dY_u^M + Y_u^M dv_u + dY_u^M dv_u\\ &= -\kappa E_t^Q\int_t^s E_t^Q(Y_u^M v_u)du + \kappa\int_t^s E_t^Q(Y_u^M\theta_u)du\\ &= -\kappa E_t^Q\int_t^s E_t^Q(Y_u^M v_u)du + \sigma_\theta^2\theta_t\int_t^s \kappa(s-u)\frac{e^{-\kappa(T-u)}-1+\kappa(T-u)}{\kappa}du.\end{aligned}$$

Solving the ordinary differential equation (ODE) gives

$$\begin{aligned}\int_t^s E_t^Q(Y_u^M v_u)du &= \sigma_\theta^2\theta_t\int_t^s e^{-\kappa(s-r)}\int_t^r \kappa(r-u)\frac{e^{-\kappa(T-u)}-1+\kappa(T-u)}{\kappa}dudr\\ &= \sigma_\theta^2\theta_t\int_t^s \frac{e^{-\kappa(s-u)}-1+\kappa(s-u)}{\kappa}\frac{e^{-\kappa(T-u)}-1+\kappa(T-u)}{\kappa}du.\end{aligned}$$

Taking differentiation with respect to s gives

$$E_t^Q(Y_s^M v_s) = \sigma_\theta^2\theta_t\int_t^s (1-e^{-\kappa(s-u)})\frac{e^{-\kappa(T-u)}-1+\kappa(T-u)}{\kappa}du,$$

which is equivalent to equation (45). This completes the proof.

B.3 Jump Diffusion Models

In a jump-diffusion model, the risk-neutral underlying stock is modeled by

$$\frac{dS_t}{S_t} = rdt + \sqrt{v_t}dB_t^S + (e^x-1)dN_t - \lambda_t(e^x-1)dt, \tag{46}$$

where r is the risk free rate, the jump size x is constant, the Brownian motion variance v_t and jump intensity λ_t are stochastic. We do not specify the processes for v_t and λ_t to derive generic expressions for the second and third central moments of logarithmic return in jump-diffusion models.

Applying Ito's Lemma to equation (46) gives

$$d\ln S_t = \Big[r - \frac{1}{2}v_t - \lambda_t(e^x - 1 - x)\Big]dt + \sqrt{v_t}dB_t^S + xdN_t - \lambda_t x dt. \tag{47}$$

The logarithmic return from current time t, to a future time, T, is defined by

$$R_t^T \equiv \ln\frac{S_T}{S_t} = \int_t^T \Big[r - \frac{1}{2}v_u - \lambda_u(e^x - 1 - x)\Big]du + \sqrt{v_u}dB_u^S + xdN_u - \lambda_u x du. \tag{48}$$

The conditional expectation at time t is then given by

$$E_t^Q(R_t^T) = \int_t^T \Big[r - \frac{1}{2}E_t^Q(v_u) - E_t^Q(\lambda_u)(e^x - 1 - x)\Big]du. \tag{49}$$

Following the notations in Zhang et al (2015), we introduce another two integrals

$$Z_T \equiv \int_t^T [\lambda_u - E_t^Q(\lambda_u)]du, \qquad I_T \equiv \int_t^T dN_u - \lambda_u du.$$

With those notations as well as X_T and Y_T, subtracting (49) from (48) yields

$$R_t^T - E_t^Q(R_t^T) = X_T - \frac{1}{2}Y_T + xI_T - (e^x - 1 - x)Z_T. \tag{50}$$

Noting that dN_u is conditionally independent of λ_u, and using the incremental independency property of a Poisson process, we obtain the variance and third cumulant of R_t^T as follows

$$\begin{aligned}
E_t^Q[R_t^T - E_t^Q(R_t^T)]^2 &= E_t^Q(X_T^2) - E_t^Q(X_TY_T) + \frac{1}{4}E_t^Q(Y_T^2) + x^2E_t^Q(I_T^2) \\
&\quad + (e^x - 1 - x)^2E_t^Q(Z_T^2), \\
E_t^Q[R_t^T - E_t^Q(R_t^T)]^3 &= E_t^Q(X_T^3) - \frac{3}{2}E_t^Q(X_T^2Y_T) + \frac{3}{4}E_t^Q(X_TY_T^2) - \frac{1}{8}E_t^Q(Y_T^3) \\
&\quad + x^3E_t^Q(I_T^3) - 3x^2(e^x - 1 - x)E_t^Q(Z_T^2) - (e^x - 1 - x)^3E_t^Q(Z_T^3),
\end{aligned}$$

where the variance of X_T, the variance and third cumulant of I_T are given by

$$E_t^Q(X_T^2) = \int_t^T E_t^Q(v_u)du, \qquad E_t^Q(I_T^2) = E_t^Q(I_T^3) = \int_t^T E_t^Q(\lambda_u)du.$$

To express the terms in variance and third cumulant of R_t^T explicitly, we need to specify the Brownian motion variance process and jump intensity process. See Zhang et al (2015) for the case of Heston model. Furthermore, if the jump intensity also follows a square root process, we obtain similar results for $E_t^Q(Z_T^2)$ and $E_t^Q(Z_T^3)$ as Heston model.

References

[1] Albuquerque, Rui, 2012, Skewness in stock returns: Reconciling the evidence on firm versus aggregate returns, *Review of Financial Studies* 25(5), 1630-1673.

[2] Bakshi, Gurdip, Nikunj Kapadia, and Dilip Madan, 2003, Stock return characteristics, skew laws, and the differential pricing of individual equity options, *Review of Financial Studies* 16(1), 101-143.

[3] Bali, Turan G., and Scott Murray, 2013, Does risk-neutral skewness predict the cross-section of equity option portfolio returns? *Journal of Financial and Quantitative Analysis* 48(4), 1145-1171.

[4] Bates, David S., 1996, Jumps and stochastic volatility: exchange rate processes implicit in deutsche mark options, *Review of Financial Studies* 9(1), 69-107.

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

[6] Black, Fischer, and Myron Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81(3), 637-654.

[7] Boyer, Brian, Todd Mitton, and Keith Vorkink, 2010, Expected idiosyncratic skewness, *Review of Financial Studies* 23(1), 169-202.

[8] Chang, Bo Young, Peter Christoffersen, and Kris Jacobs, 2013, Market skewness risk and the cross section of stock returns, *Journal of Financial Economics* 107(1), 46-68.

[9] Chen, Joseph, Harrison Hong, and Jeremy C Stein, 2001, Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices, *Review of Financial Studies* 61(3), 345-381.

[10] Conrad, Jennifer, Robert F. Dittmar, and Eric Ghysels, 2013, Ex ante skewness and expected stock returns, *Journal of Finance* 68(1), 85-124.

[11] Das, Sanjiv Ranjan, and Rangarajan K. Sundaram, 1999, Of smiles and smirks: A term-structure perspective, *Journal of Financial and Quantitative Analysis* 34(2), 211-240.

[12] DeMiguel, Victor, Yuliya Plyakha, Raman Uppal, and Grigory Vilkov, 2013, Improveing portfolio selection using option-implied volatility and skewness, *Journal of Financial and Quantitative Analysis* 48(6), 1813-1845.

[13] Dennis, Patrick, and Stewart Mayhew, 2002, Risk-neutral skewness: Evidence from stock options, *Journal of Financial and Quantitative Analysis* 37(3), 471-493.

[14] Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* 68(6), 1343-1376.

[15] Eraker, Bjørn, Michael Johannes, and Nicholas Polson, 2003, The impact of jumps in volatility and returns, *Journal of Finance* 58(3), 1269-1300.

[16] Faff, Robert W., and Zhangxin Liu, 2014, Hitting SKEW for SIX, Working paper, University of Western Australia.

[17] Han, Bing, 2008, Investor sentiment and option prices, *Review of Financial Studies* 21(1), 387-414.

[18] Harvey, Campbell R., and Akhtar Siddique, 2000, Conditional skewness in asset pricing tests, *Journal of Finance* 55(3), 1263-1295.

[19] Heston, Steven L., 1993, A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* 6(2), 327-343.

[20] Kraus, Alan, and Robert H. Litzenberger, 1976, Skewness preference and the valuation of risk assets, *Journal of Finance* 31(4), 1085-1100.

[21] Luo, Xingguo, and Jin E. Zhang, 2012, The term structure of VIX, *Journal of Futures Markets* 32(12), 1092-1123.

[22] Merton, Robert C., 1976, Option pricing when underlying stock returns are discontinuous, *Journal of Financial Economics* 3(1-2), 125-144.

[23] Mitton, Todd, and Keith Vorkink, 2007, Equilibrium underdiversification and the preference for skewness, *Review of Financial Studies* 20(4), 1255-1288.

[24] Neuberger, Anthony, 2012, Realized skewness, *Review of Financial Studies* 25(11), 3423-3455.

[25] Wang, Zhiguang, and Robert T. Daigler, 2012, The option SKEW index and the volatility of volatility, Working paper, South Dakota State University.

[26] Zhang, Jin E., Huimin Zhao, and Eric C. Chang, 2012, Equilibrium asset and option pricing under jump diffusion, *Mathematical Finance* 22(3), 538-568.

[27] Zhang, Jin E., Fang Zhen, Xiaoxia Sun, and Huimin Zhao, 2015, The skewness implied in Heston model, Working paper, University of Otago.

[28] Zhao, Huimin, Jin E. Zhang, and Eric C. Chang, 2013, The Relation between physical and risk-neutral cumulants, *International Review of Finance* 13(3), 345-381.

Table 1: Descriptive statistics of daily SKEW term structure

We interpolate the daily SKEW term structure to constant time to maturity, $\tau = 1$, 3, 6, 9, 12 and 15 months. The sample period is from 2 January 1990 to 31 December 2014. We only consider the days when the maximum time to maturity is not less than 15 months.

Maturity	Mean	Std Dev	Skewness	Kurtosis	Minimum	Maximum
1 Month	117.8912	6.1185	0.7931	3.9948	101.0830	147.0000
3 Months	116.7662	5.3919	0.6332	5.8006	96.8083	173.2275
6 Months	116.5344	5.6636	0.7390	3.7362	99.5674	151.1067
9 Months	115.0219	6.4143	0.8855	4.1927	99.2181	165.8313
12 Months	114.5970	6.9751	0.8138	3.4740	97.0592	147.2174
15 Months	114.6724	7.5005	1.0559	4.0158	96.4798	155.6445

Figure 1: The time evolution of the SKEW term structure.

This graph shows the time evolution of the interpolated SKEW term structure for time to maturity from 1 month to 15 months. The sample period is from 2 January 1990 to 31 December 2014. We only consider the days when the maximum time to maturity is not less than 15 months.

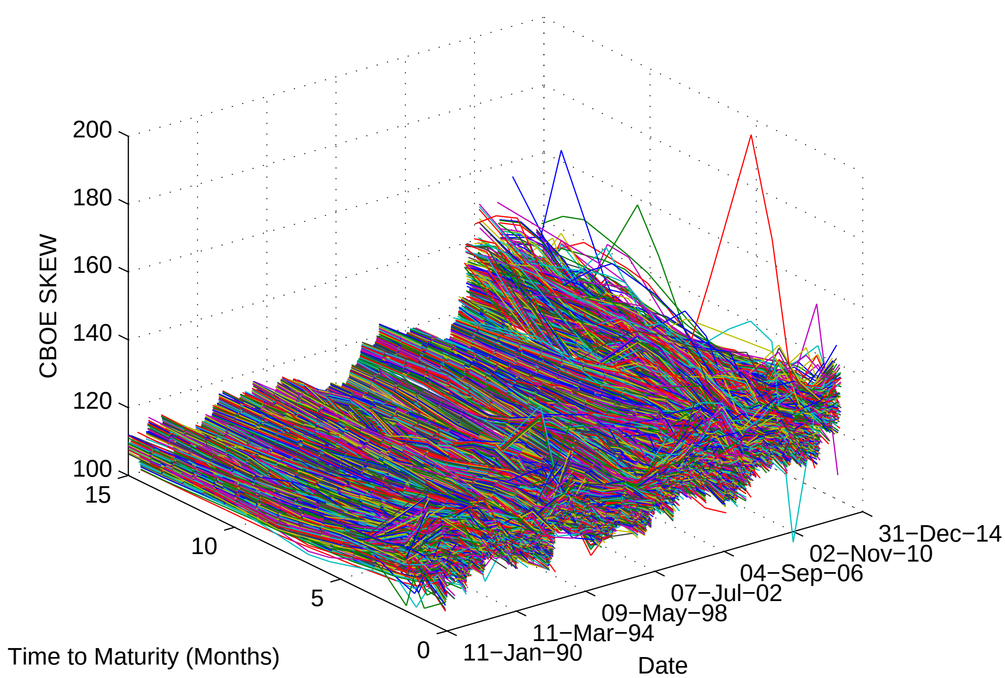

Figure 2: The time series of SKEW with constant time to maturity
This graph shows the time series of SKEW with 1 month, 6 months and 15 months to maturity. The sample period is from 2 January 1990 to 31 December 2014. We only consider the days when the maximum time to maturity is not less than 15 months.

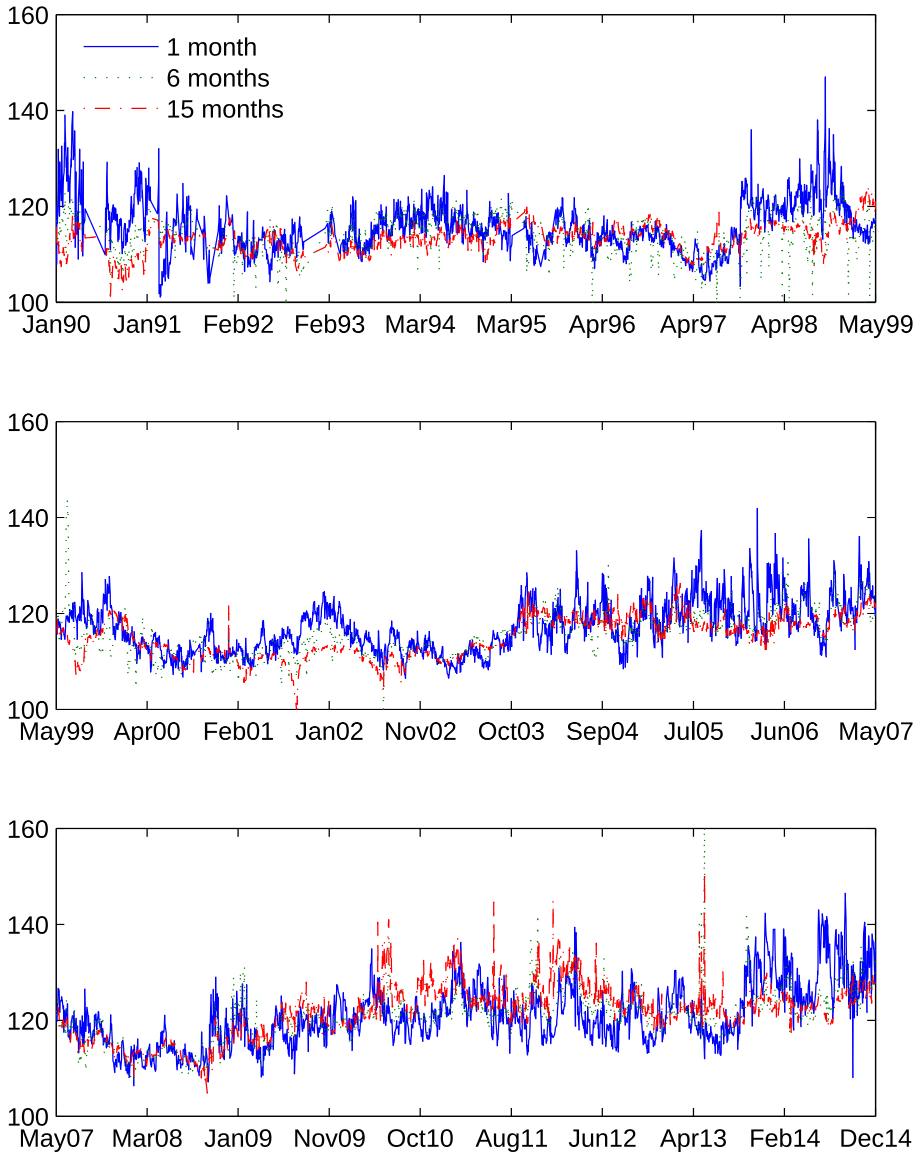

Figure 3: A few samples of SKEW term structures with abnormal shape.

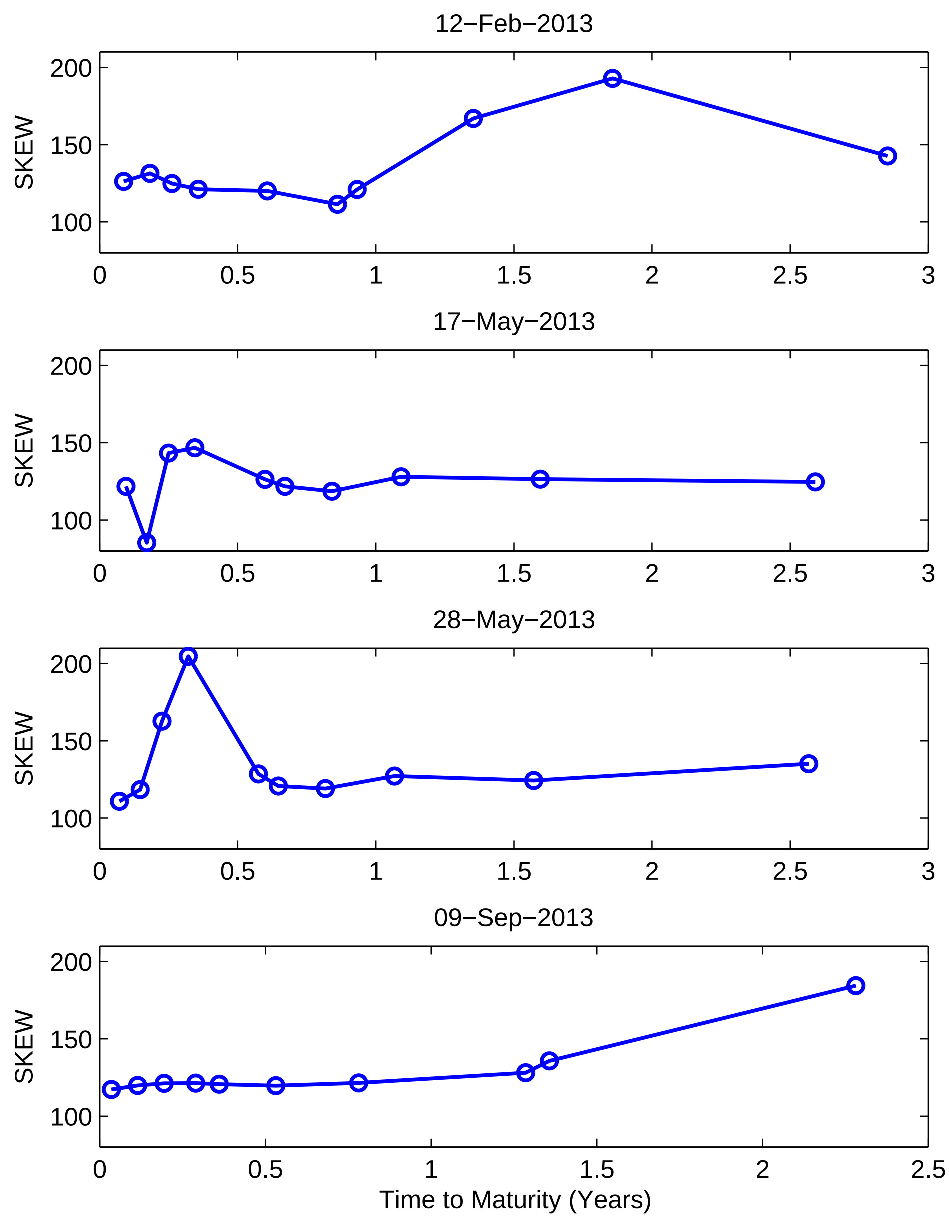

Table 2: Optimal Values for the Mean-reverting Speed.
Fitting the VIX term structure data of different periods, we obtain the optimal values for the mean-reverting speed κ in terms of minimum root mean squared error (RMSE), as well as the time series of instantaneous squared VIX V_t and its long term mean level θ_t^V. $\overline{V}_t$ and $\overline{\theta}_t$ denote the means of V_t and θ_t^V. $\sigma(V_t)$ and $\sigma(\theta_t)$ denote the standard deviations of V_t and θ_t^V. We select the days with no less than 6 maturities for both VIX and SKEW to guarantee the estimation precision.

κ	RMSE	$\overline{V}_t$	$\sigma(V_t)$	$\overline{\theta}_t$	$\sigma(\theta_t)$
Panel A: 2 January 1992 to 4 September 2009					
1	1.163	0.0467	0.0486	0.0419	0.0271
2	1.057	0.0472	0.0529	0.0429	0.0258
3	1.012	0.0474	0.0556	0.0435	0.0266
4	0.9891	0.0476	0.0581	0.0439	0.0274
5	0.9795	0.0477	0.0603	0.0441	0.0280
5.8285	0.9776	0.0478	0.0621	0.0442	0.0284
6	0.9776	0.0478	0.0624	0.0442	0.0285
7	0.9804	0.0479	0.0645	0.0443	0.0289
Panel B: 24 November 2010 to 31 December 2014					
1	0.6482	0.0330	0.0294	0.0902	0.0295
1.9304	0.5622	0.0310	0.0303	0.0744	0.0262
2	0.5626	0.0308	0.0303	0.0738	0.0261
3	0.6240	0.0290	0.0312	0.0680	0.0254
4	0.7283	0.0275	0.0321	0.0649	0.0252
5	0.8343	0.0262	0.0329	0.0628	0.0250
6	0.9314	0.0250	0.0337	0.0614	0.0250
7	1.018	0.0239	0.0345	0.0602	0.0249
Panel C: Full Sample					
1	1.071	0.0441	0.0459	0.0512	0.0335
2	0.9693	0.0441	0.0498	0.0488	0.0286
3	0.9396	0.0439	0.0523	0.0482	0.0281
3.5680	0.9361	0.0438	0.0536	0.0480	0.0281
4	0.9375	0.0437	0.0546	0.0479	0.0282
5	0.9493	0.0436	0.0567	0.0477	0.0284
6	0.9676	0.0434	0.0587	0.0475	0.0287
7	0.9889	0.0433	0.0606	0.0474	0.0289

Table 3: Optimal Value for the Volatility of Volatility.
Set jump size $x = -0.14$, correlation coefficient $\rho = -0.859$, given the optimal time series V_t and θ_t^V obtained from the full sample VIX term structure, fitting the SKEW term structure, we get the optimal value of σ_v 0.7929 in terms of minimum root mean squared error (RMSE), as well as the time series of Brownian variance v_t and its long term mean level θ_t^v. $\overline{v}_t$ and $\overline{\theta_t^v}$ denote the means of v_t and θ_t^v. $\sigma(v_t)$ and $\sigma(\theta_t^v)$ denote the standard deviations of v_t and θ_t^v. We select the days with no less than 6 maturities for both VIX and SKEW to guarantee the estimation precision.

σ_v	RMSE	$\overline{v}_t$	$\sigma(v_t)$	$\overline{\theta_t^v}$	$\sigma(\theta_t^v)$
0.6	7.5287	0.0265	0.0374	0.0330	0.0350
0.7	7.2350	0.0287	0.0436	0.0292	0.0353
0.7929	7.1561	0.0284	0.0458	0.0263	0.0350
0.8	7.1565	0.0283	0.0459	0.0261	0.0350
0.9	7.2382	0.0267	0.0471	0.0238	0.0341
1.0	7.4201	0.0247	0.0475	0.0217	0.0324

Figure 4: The estimate time series of instantaneous squared VIX.
This graph shows the daily optimal values of instantaneous squared VIX and its long term mean level with the sample period from 2 January 1992 to 4 September 2009 and 24 November 2010 to 31 December 2014. The statistics of the two series are shown in Table 2 with optimal mean-reverting speed κ 3.5680.

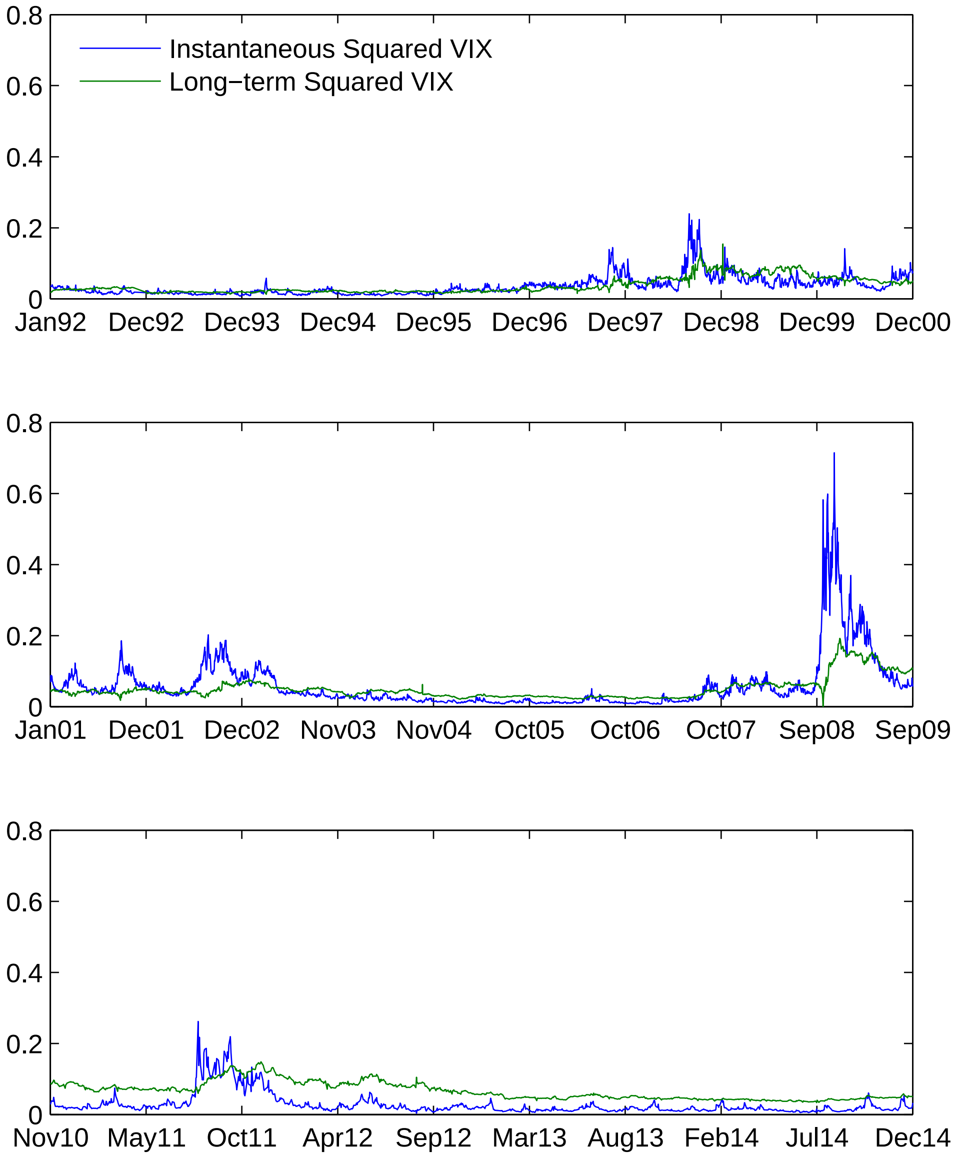

Figure 5: The estimate time series of Brownian variance.

This graph shows the daily optimal values of Brownian variance and its long term mean level with the sample period from 2 January 1992 to 4 September 2009 and 24 November 2010 to 31 December 2014. The statistics of the two series are shown in Table 3 with optimal volatility of volatility σ_v 0.7929.

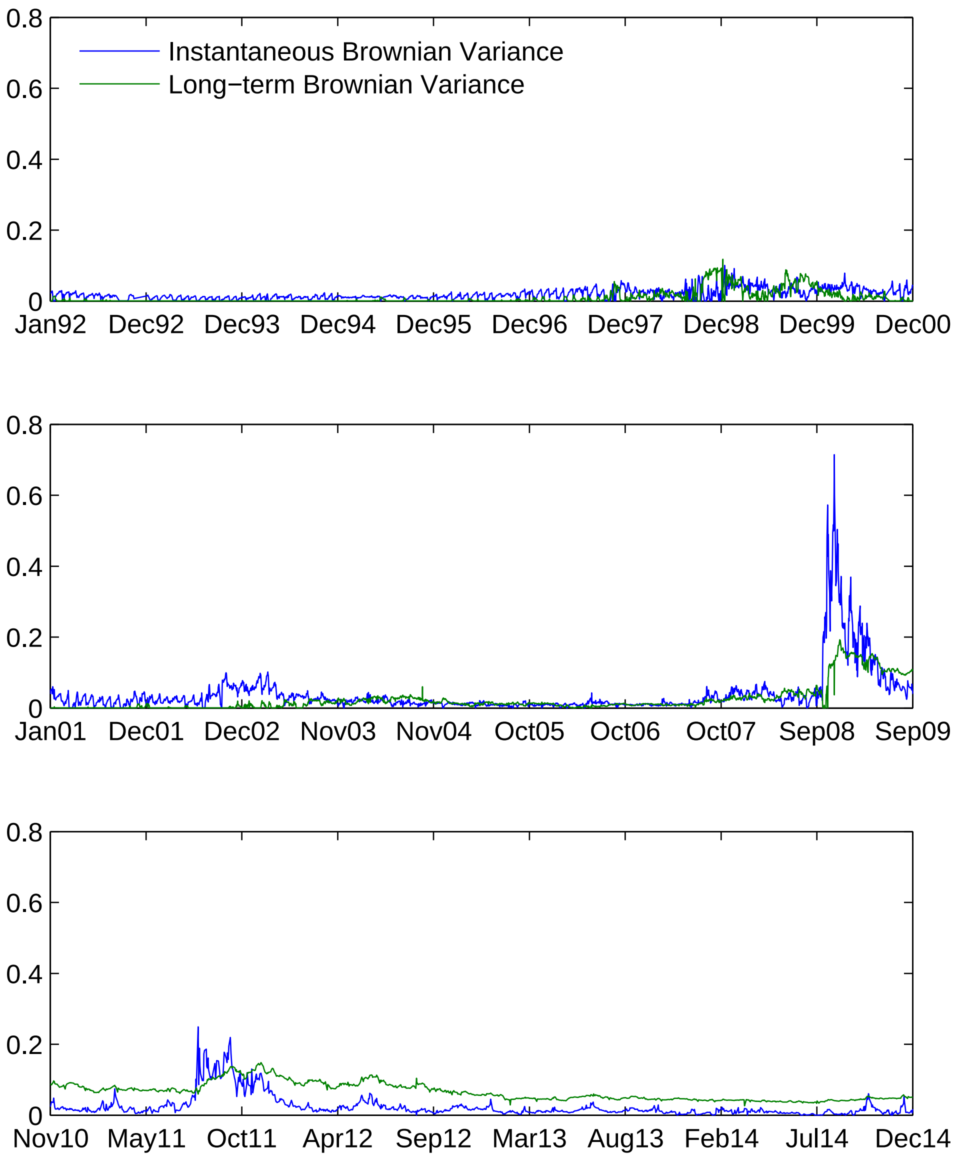

Figure 6: The estimate time series of jump intensity.
This graph shows the time series of jump intensity and its long term mean level, which can be directly obtained from the daily squared VIX and Brownian variance with the sample period from 2 January 1992 to 4 September 2009 and 24 November 2010 to 31 December 2014. The average of spot jump intensity is 0.818, whereas the average of long-term jump intensity is 1.153.

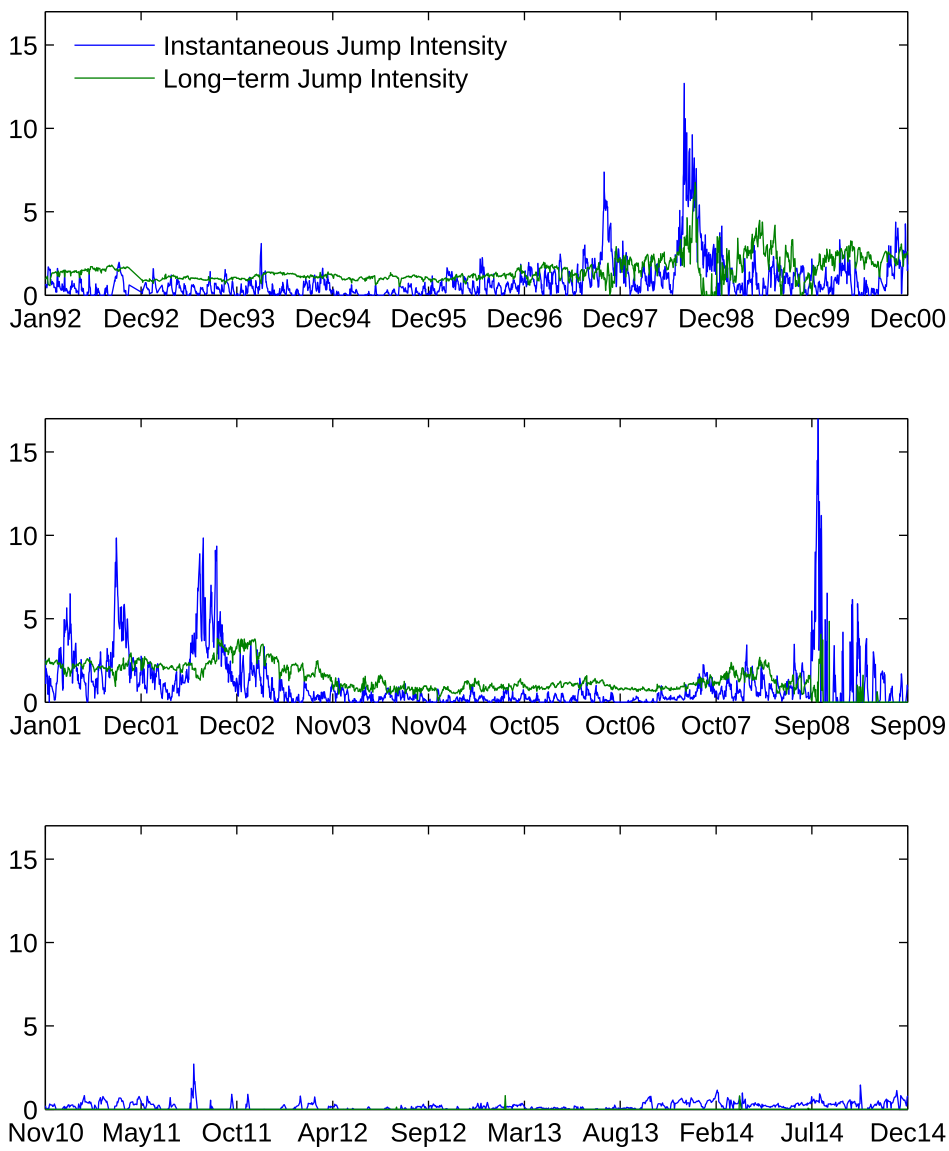

Figure 7: The time series of instantaneous variance ratio and regularized skewness. The average of instantaneous Brownian variance to squared VIX is 66.9%, and the average of regularized skewness at point zero is -0.27.

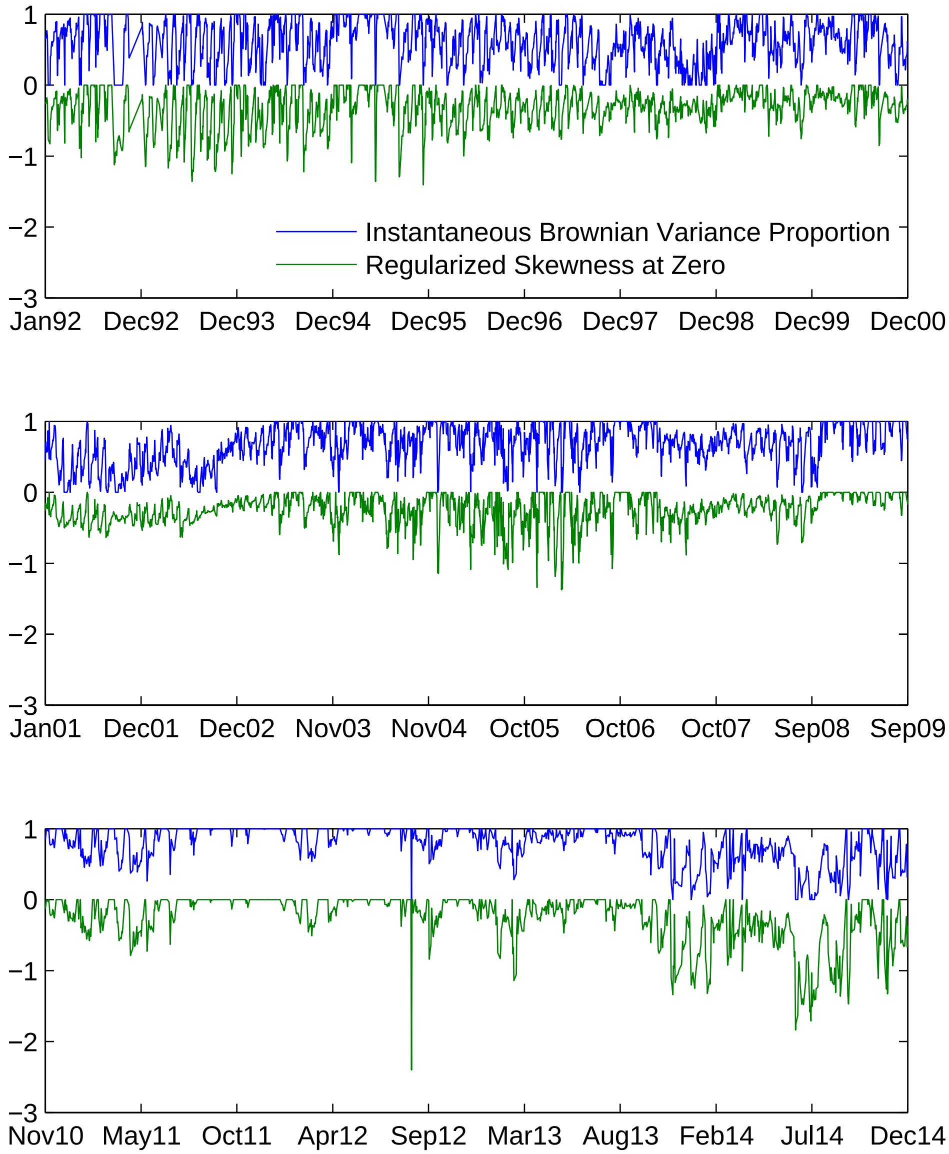

Figure 8: Nonlinearity between Brownian Variance and Jump Intensity
This graph shows the scatter of the Brownian variance and jump intensity estimated on the same day. The correlation coefficient between the two time series is 0.055.

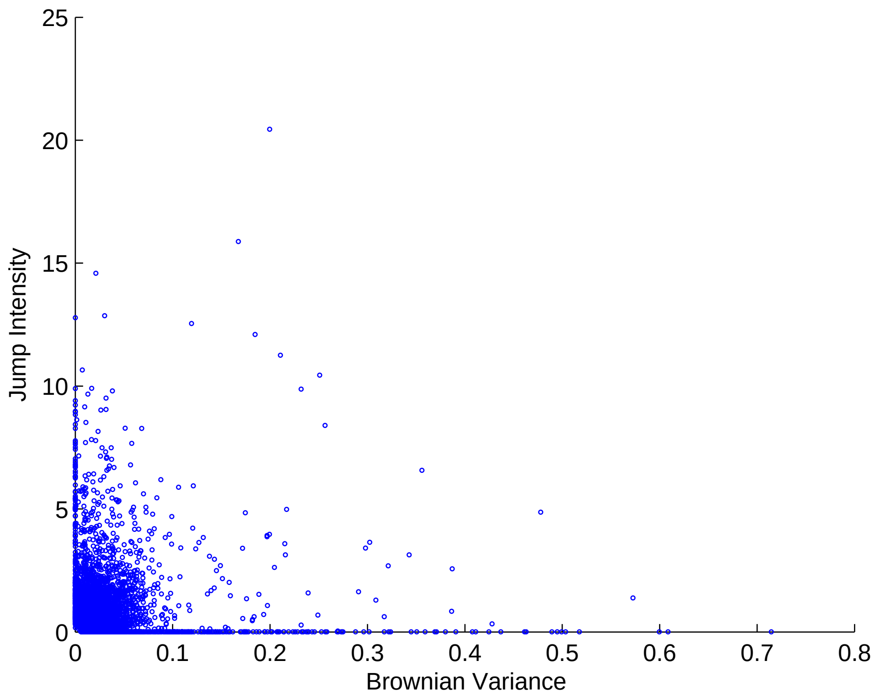

Figure 9: Model-implied SKEW.

We use the optimal average values of V_t (0.0438), θ_t^V (0.048) derived from VIX term structure estimation, θ_t^v (0.0263) derived from SKEW term structure estimation, and the Brownian variance v_t (0.0311), jump size x (-0.14), correlation coefficient ρ (-0.859) derived from the assumption that Brownian variance is 70% of total variance on average as the inputs of the model-implied SKEW. The red dots represent the weekly SKEW averages with 2 to 52 weeks to maturity, whereas the green dots represent the monthly interpolated SKEW with 1 to 12 months to maturity.

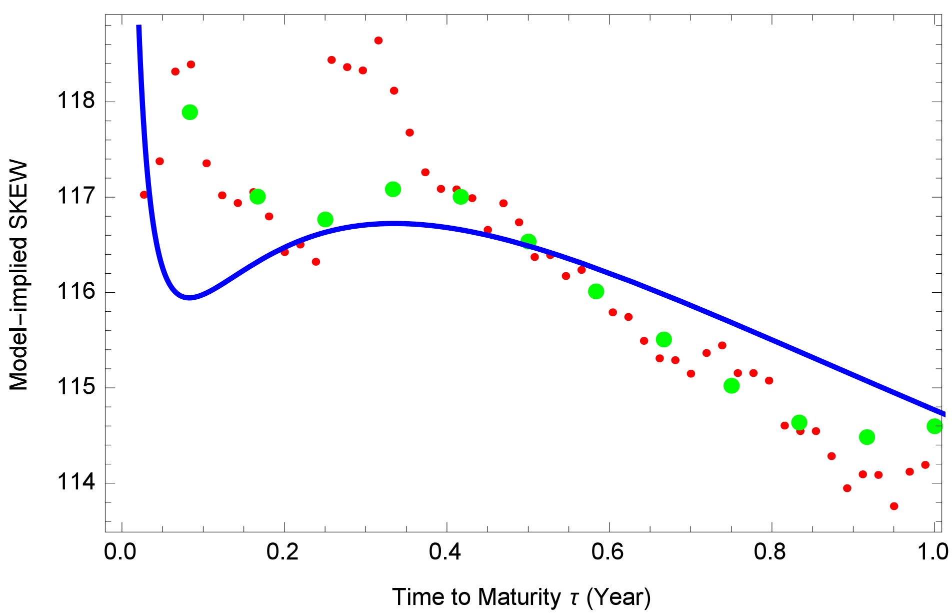

Figure 10: Model-implied Regularized Skewness.
The blue line represents the model-implied regularized skewness, whose inputs are the optimal average values of V_t (0.0438), θ_t^V (0.048) derived from VIX term structure estimation, the Brownian variance v_t (0.0311), jump size x (-0.14), correlation coefficient ρ (-0.859) derived from the assumption that Brownian variance is 70% of total variance on average, and the optimal average value of θ_t^v (0.0263) derived from SKEW term structure estimation. The dashed lines represent the asymptote of the model-implied regularized skewness at zero and infinity. The dots represent the weekly averages, from 2 weeks to 52 weeks, of regularized skewness, which is defined as the product of the true skewness and the square root of time to maturity.