Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

The Effects of Asymmetric Volatility and Jumps on the Pricing of VIX Derivatives

Yang-Ho Park

2015-071

Please cite this paper as:
Park, Yang-Ho (2015). "The Effects of Asymmetric Volatility and Jumps on the Pricing of VIX Derivatives," Finance and Economics Discussion Series 2015-071. Washington: Board of Governors of the Federal Reserve System, http://dx.doi.org/10.17016/FEDS.2015.071.

The Effects of Asymmetric Volatility and Jumps on the Pricing of VIX Derivatives

Yang-Ho Park

Federal Reserve Board

September 11, 2015

ABSTRACT

This paper proposes a new collection of affine jump-diffusion models for the valuation of VIX derivatives. The models have two distinctive features. First, we allow for a positive correlation between changes in the VIX and in its stochastic volatility to accommodate asymmetric volatility. Second, upward and downward jumps in the VIX are separately modeled to accommodate the possibility that investors react differently to good and bad surprises. Using the VIX futures and options data from July 2006 through January 2013, we find conclusive evidence for the benefits of including both asymmetric volatility and upward jumps in models of VIX derivatives pricing. We do not, however, find evidence supporting downward jumps.

JEL Classification: G12; G13

Keywords: VIX options; VIX futures; stochastic volatility; volatility smile; jump-diffusion

We are very grateful for comments and suggestions from Yacine Aït-Sahalia, Michael Gordy, Jaideep Oberoi, and seminar participants at the IFSID Third Conference on Derivatives. This paper benefits from the excellent research assistance of Juliette Lu. Disclaimer: The analysis and conclusions set forth are those of the authors and do not indicate concurrence by the Board of Governors or other members of its research staff. Send correspondence to Yang-Ho Park, Risk Analysis Section, Board of Governors of the Federal Reserve System, 20th & C Streets, NW, Washington, D.C. 20551, Tel: (202) 452-3177; e-mail: yang-ho.park@frb.gov.

1 Introduction

Over the past few years, there has been a great deal of financial innovation in volatility trading markets. A new collection of volatility derivatives, such as VIX futures, options, and exchange-traded products, have been introduced, making volatility trading more accessible to a broad range of investors.[1] The increasing volume of trading in those products is largely due to the fact that they can be used as a so-called tail risk hedging strategy against stock market downturns. That is, because changes in market volatility are negatively correlated with stock market returns, investors can limit the loss of an equity portfolio by taking a long position in VIX futures or call options. Given the explosive growth in volatility trading, the objective of this paper is to understand the effects of asymmetric volatility of and jumps in the VIX on the pricing of VIX futures and options.

We use the term "asymmetric volatility" to refer to the fact that the volatility of the VIX is not merely stochastic but also varies asymmetrically in response to changes in the VIX. Such asymmetry can be seen, for example, by looking at the dynamic relation between the VIX and the VVIX of the CBOE (Chicago Board Options Exchange). The VIX is a risk-neutral, forward-looking measure of market volatility implied by a cross section of S&P 500 index options, while the VVIX is a risk-neutral, forward-looking measure of market volatility of volatility implied by a cross section of VIX options.[2] The scatter plot in Figure 1 shows that there is a positive relation between changes in the VVIX and in the VIX. That is, the volatility of the VIX, as measured by the VVIX, tends to increase (decrease) as the VIX increases (decreases).

An important implication of this empirical regularity is that the stochastic volatility factor implicit in the VIX options comprises two components—one that can be spanned by VIX futures and another that cannot. In other words, the VIX options market is nonredundant with the VIX futures market. To gauge the extent to which the options market can be spanned by the futures market, we run a regression of VVIX returns onto VIX returns, and find that the VIX changes can explain only 51 percent of the variation in the VVIX based on the adjusted R^2 of the regression.

[1] See, for example, Whaley (2013) and Alexander and Korovilas (2012) as references on exchange-traded volatility products linked to the VIX.

[2] The VVIX is computed by applying the VIX formula to the VIX options market.

Contrary to the empirical fact just discussed, the existing VIX option papers generally assume that the volatility of the VIX is either completely spanned or completely unspanned. For example, some authors, such as Grünbichler and Longstaff (1996), Detemple and Osakwe (2000), and Goard and Mazur (2013), apply fully spanned volatility models in which VIX options are completely hedgeable by VIX futures, whereas others, such as Mencía and Sentana (2012), introduce a completely unspanned volatility model in which innovations in the VIX index are uncorrelated with those in the volatility of the VIX index. It is evident that all of those models are unable to accommodate the important feature of the data that we have found—namely, asymmetric volatility. Hence, the chief goal of this paper is to contribute to the literature by studying the effects of asymmetric volatility on the pricing of VIX derivatives.

Intuitively, a model that allows for asymmetry in volatility may have great potential to improve the pricing of VIX options because it is capable of explaining the positive skewness implicit in the options.[3] The VIX options market shows a persistent deviation from a geometric Brownian motion regardless of time to maturity. In particular, out-of-the-money (OTM) VIX calls tend to have higher Black-Scholes implied volatility than OTM VIX puts, a tendency that is sometimes referred to as a volatility smile, implying that the option-implied VIX return distributions are positively skewed. This positive skewness may be in part attributable to asymmetric volatility.

On top of asymmetric volatility, we also look at the impact of jumps in the VIX on the valuation of VIX futures and options because they can be another channel for explaining the positive skewness implied by VIX options. Good and bad surprises may arrive with different rates and sizes and investors may react differently to them.[4] Hence, our models assume that upward and downward jumps occur independently with different frequencies and magnitudes. In this respect, our jump treatment is close in spirit to the currency option pricing model of Carr and Wu (2007), which separately models upward and downward jumps using time-changed Levy processes.

[3] This echoes a well-known finding that the negative skewness implicit in stock index options is partly associated with the leverage or volatility feedback effect, which indicates a negative correlation between stock index returns and volatility changes.

[4] The market's asymmetric response to shocks or news announcements has been documented by Andersen, Bollerslev, Diebold, and Vega (2007) and Bakshi, Carr, and Wu (2008), among others.

Building on the affine jump-diffusion framework of Duffie, Pan, and Singleton (2000), we introduce a new family of dynamic models for the VIX with and without the features of asymmetric volatility and jumps and derive quasi-analytic solutions to the prices of futures and options. The models are characterized by three dynamic factors: one observed VIX index and two latent factors to capture time variations in the stochastic volatility and central tendency of the VIX. To reflect asymmetric volatility, we allow for a nonzero correlation between innovations in the VIX and in its stochastic volatility. Upward and downward jumps are both assumed to follow independent compound Poisson processes with each having its own jump intensity and exponential jump-size distribution. The models are then tested on the VIX futures and options data covering July 2006 through January 2013, via an unscented Kalman filter.

Turning to the empirical results, we first evaluate the effects of asymmetric volatility on the pricing of VIX derivatives. By comparing the model with symmetric volatility and no jumps (SVV) with the model with asymmetric volatility and no jumps (AVV), we find that the latter is strongly preferred to the former in both in-sample and out-of-sample tests. The performance difference is statistically significant at the 1 percent level based on the Diebold and Mariano (2002) test. That is, allowing for asymmetry in volatility can make large improvements in fitting the prices of VIX futures and options.

We next look at the effects of including upward jumps on top of asymmetric volatility by comparing the AVV model with the AVV-UJ model, the model with asymmetric volatility and upward jumps. The comparison shows decisive evidence in favor of the AVV-UJ model over the AVV model, except for in-sample futures pricing. The results are statistically significant at the 1 percent level according to the Diebold and Mariano (2002) test. That is, including upward jumps in the AVV model can make considerable improvements in VIX derivatives pricing.

Lastly, we investigate whether downward jumps can have an incremental effect on the pricing of VIX derivatives by comparing the AVV-UJ model with the model with asymmetric volatility and asymmetric upward and downward jumps (AVV-AJ). The comparison does not result in any significant ranking of one model over the other. Including downward jumps makes little, if any, improvement in the pricing of both futures and options as long as asymmetric volatility and upward jumps are accounted

for. However, because we consider finite-activity jumps in this paper, we do not interpret our finding as implying that downward jumps are of no use in explaining VIX derivatives prices; rather, we argue that upward jumps play a far more important role in the pricing of VIX derivatives than downward jumps do.

Aside from the VIX option pricing literature, this paper also adds to the literature on unspanned stochastic volatility. Unspanned volatility has been found to be important in many other asset classes. For example, unspanned volatility in the interest rate market has been reported by Collin-Dufresne and Goldstein (2002), Li and Zhao (2006), and Bikbov and Chernov (2009), among others. Trolle and Schwartz (2009b) emphasize the importance of including unspanned volatility in pricing commodity options.

This paper is also related to the literature on the structure of jumps in the VIX. For example, Todorov and Tauchen (2011) and Wu (2011) both find that jumps in the VIX are more likely to have infinite activity, although they rely on different sources of information and estimation. More recently, Todorov, Tauchen, and Grynkiv (2014) examine asymmetry in upward and downward jumps in VIX dynamics and find that upward and downward jumps are approximately symmetric, although the former are slightly more active than the latter. In more loosely related papers, Eraker (2005) and Broadie, Chernov, and Johannes (2007) study the impact of volatility jumps in joint analyses of stock index returns and options data. Our paper is distinct from all of these papers in that our application takes a direct look at VIX options and futures, which may offer a better environment in which to scrutinize the characteristic of volatility jumps.

The rest of the paper is organized as follows. Section 2 develops the models and methods used in this paper. The VIX options and futures data are described in Section 3. Section 4 shows the parameter estimates and some preliminary analysis, while Section 5 provides the main empirical results on pricing performance across the different model specifications. Section 6 concludes.

2 Models and Methods

The affine jump-diffusion framework of Duffie, Pan, and Singleton (2000) has been widely used in the term structure modeling of interest rates and the valuation of

derivatives because it allows for analytic tractability. Dai and Singleton (2000); Heston (1993); Bakshi, Cao, and Chen (1997); and Bates (2000) are well-known examples. In this section, we introduce a new family of affine jump-diffusion models for the dynamics of the logarithm of the VIX and derive semi-closed-form solutions to the prices of futures and options. In addition, we provide a description of our estimation method and introduce a competing model from Mencía and Sentana (2012), which will be used as a benchmark.

2.1 VIX dynamics

We assume that the logarithm of the VIX follows an affine jump-diffusion process. This assumption is motivated by a well-known empirical finding that a logarithmic model does a better job of describing the volatility dynamics of a stock index than the square root model of Heston (1993) (see, for example, Aït-Sahalia and Kimmel (2007), Jones (2003), and Durham (2013)). More importantly, Mencía and Sentana (2012) compare logarithmic and square root models with respect to the pricing of VIX futures and options and find that the former is preferred to the latter. Furthermore, an affine modeling of the logarithm of the VIX is important because it permits negative jumps as well as positive ones, whereas an affine modeling of the VIX itself leaves no room for modeling negative jumps.

As stated earlier, the objective of this paper is to understand the effects of asymmetric volatility of and jumps in the VIX on the pricing of VIX derivatives. Toward this end, it is desirable to start with a fairly realistic baseline model upon which new model characteristics are built. Our baseline model contains two latent factors in addition to the logarithmic VIX index, denoted by $v_t = \log(\text{VIX}_t)$. Specifically, a latent factor, u_t, is introduced to reflect time variation in the long-run mean of the VIX, which greatly helps to match VIX future prices along increasing times to maturity.[5] The volatility of the VIX is also allowed to vary over time, driven by another latent factor, w_t.

Given a risk-neutral probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ and information filtration $\{\mathcal{F}_t\}$,

[5] See Gallant, Hsu, and Tauchen (1999), Christoffersen, Heston, and Jacobs (2009), and Aït-Sahalia, Amengual, and Manresa (2015), among others, for the importance of a long-run volatility factor.

we assume that the (logarithmic) VIX dynamics take the following form:

$$\begin{aligned} dv_t &= \kappa_v(u_t - v_t)dt + \sqrt{w_t}dB_{1t}^{\mathbb{Q}} + J_1^{\mathbb{Q}}dN_{1t}^{\mathbb{Q}} + J_2^{\mathbb{Q}}dN_{2t}^{\mathbb{Q}} - \lambda_+\delta_+dt - \lambda_-\delta_-dt \\ du_t &= \kappa_u(\overline{u} - u_t)dt + \sigma_u dB_{2t}^{\mathbb{Q}} \\ dw_t &= \kappa_w(\overline{w} - w_t)dt + \sigma_w\sqrt{w_t}dB_{3t}^{\mathbb{Q}}, \end{aligned} \tag{1}$$

where κ_v, κ_u, and κ_w capture the persistence of v_t, u_t, and w_t, respectively; σ_u and σ_w control for the volatility of u_t and w_t, respectively; $\overline{u}$ and $\overline{w}$ capture the long-run means of u_t and w_t, respectively; and $B_{1t}^{\mathbb{Q}}$, $B_{2t}^{\mathbb{Q}}$, and $B_{3t}^{\mathbb{Q}}$ are standard Brownian motions under the risk-neutral $\mathbb{Q}$ measure.

To capture the asymmetric behavior of volatility, we allow for a nonzero correlation, denoted by ρ, between $dB_{1t}^{\mathbb{Q}}$ and $dB_{3t}^{\mathbb{Q}}$. Although we expect ρ to be positive, we do not restrict the sign of ρ, letting the data speak as to its direction.

We allow for jumps in v_t. Specifically, as one of our main research goals is to understand the differential pricing implications of upward and downward jumps, we assume that they are driven by independent compound Poisson processes, with each having its own jump intensity and jump-size distribution. $N_{1t}^{\mathbb{Q}}$ and $N_{2t}^{\mathbb{Q}}$ denote risk-neutral Poisson processes driving upward and downward jumps with jump intensities λ_+ and λ_-, respectively. Upward jump magnitudes, $J_1^{\mathbb{Q}}$, are assumed to follow an independent exponential distribution with a positive mean, $\delta_+ > 0$, with the probability density function taking $\frac{1}{\delta_+}\exp(-x/\delta_+)$ if $x > 0$ and 0 otherwise. Similarly, downward jump magnitudes, $J_2^{\mathbb{Q}}$, are assumed to follow an independent exponential distribution with a negative mean, $\delta_- < 0$, with the probability density function taking $\frac{1}{|\delta_-|}\exp(-x/\delta_-)$ if $x < 0$ and 0 otherwise.

An affine jump-diffusion model is a Markov process whose drift vector, variance matrix, and jump intensities all have affine dependence on the state vector (see Duffie, Pan, and Singleton (2000)). To satisfy this condition, we assume zero correlations between $dB_{1t}^{\mathbb{Q}}$ and $dB_{2t}^{\mathbb{Q}}$ and between $dB_{2t}^{\mathbb{Q}}$ and $dB_{3t}^{\mathbb{Q}}$.[6]

The general form described in Equation (1) nests a range of model specifications with different features. Among those, we select four different specifications depending on whether there are restrictions on ρ, λ_+, or λ_-. The model specifications considered in this paper are summarized in Table 1.

[6] Otherwise the conditional variance matrix of the model in Equation (1) will lose affine dependence on the state vector.

2.2 Characteristic function

The virtue of affine models is that various transforms, such as a characteristic function, can be quasi-analytically calculated by solving a system of Riccati ordinary differential equations (ODEs). Specifically, given the model in Equation (1), we consider the characteristic function of v_T at time t under the $\mathbb{Q}$ measure, which is denoted by $g(v_t, u_t, w_t, t, T; \phi)$:

$$g(v_t, u_t, w_t, t, T; \phi) = E_t^{\mathbb{Q}} \left[\exp(\mathrm{i}\phi v_T)\right], \tag{2}$$

where $E_t^{\mathbb{Q}}$ denotes the conditional expectation using the information up to time t under the $\mathbb{Q}$ measure and $\mathrm{i} = \sqrt{-1}$.

Proposition 1. *Under some regularity conditions, the characteristic function takes an exponentially affine form of* (v_t, u_t, w_t)*:*

$$g(v_t, u_t, w_t, t, T; \phi) = \exp\left[\alpha(s) + \beta_v(s)v_t + \beta_u(s)u_t + \beta_w(s)w_t\right], \tag{3}$$

where $s = T - t$ *and the coefficients,* $\alpha(s)$*,* $\beta_v(s)$*,* $\beta_u(s)$*, and* $\beta_w(s)$*, satisfy the system of ODEs:*

$$\begin{aligned}
\dot{\alpha}(s) &= \kappa_u \overline{u} \beta_u(s) + \kappa_w \overline{w} \beta_w(s) + \frac{1}{2}\sigma_u^2 \beta_u(s)^2 \\
&\quad - [\lambda_+ \delta_+ + \lambda_- \delta_-]\beta_v(s) + \lambda_+ \left[\frac{1}{1 - \delta_+ \beta_v(s)} - 1\right] + \lambda_- \left[\frac{1}{1 - \delta_- \beta_v(s)} - 1\right] \\
\dot{\beta}_v(s) &= -\kappa_v \beta_v(s) \\
\dot{\beta}_u(s) &= \kappa_v \beta_v(s) - \kappa_u \beta_u(s) \\
\dot{\beta_w}(s) &= -\kappa_w \beta_w(s) + \frac{1}{2}\beta_v(s)^2 + \frac{1}{2}\sigma_w^2 \beta_w(s)^2 + \rho \sigma_w \beta_v(s) \beta_w(s),
\end{aligned} \tag{4}$$

where the boundary conditions are given as $\alpha(0) = 0$*,* $\beta_v(0) = \mathrm{i}\phi$*,* $\beta_u(0) = 0$*, and* $\beta_w(0) = 0$*.*

Proof. See Appendix A. □

The system of ODEs in Equation (4) can be solved by numerical solvers such as the Runge-Kutta method. Using this method, however, raises a numerical issue. When ODEs are stiff, a numerical solver requires a smaller step size in order to achieve absolute stability, impairing computational efficiency (see, for example, Huang and Yu (2007)). The stiffness of a system of ODEs is associated with how much the

eigenvalues of the system differ in orders of magnitude; the larger the difference, the stiffer the system. The eigenvalues of our system of ODEs are equal to the persistence parameters, κ_v, κ_u, and κ_w. Among the parameters, κ_v and κ_u have a wide discrepancy in orders of magnitude, as can be seen in Table 3. Specifically, κ_v is roughly 20 to 25 times as large as κ_u. Because of this stiffness in our system, it is desirable to use an implicit method instead of an explicit method, and we find that the Matlab `ode15s` solver works well in our application.[7]

2.3 Valuation of VIX futures and options

Once the characteristic function has been calculated, it is straightforward to compute the prices of futures and options. Let $F(t,T)$ denote the time-t price of a futures contract with a maturity of T. $F(t,T)$ represents the risk-neutral expectation of the time-T VIX index given the information at time t.

Proposition 2. *$F(t,T)$ is given by*

$$F(t,T) = g(v_t, u_t, w_t, t, T; \phi = -\mathrm{i}). \tag{5}$$

Proof. See Appendix A. □

Let $C(t,T,K)$ and $P(t,T,K)$ denote the time-t prices of call and put options, respectively, with a maturity of T and a strike price of K. These prices are determined by

$$\begin{aligned} C(t,T,K) &= \exp(-r_t\tau)E_t^{\mathbb{Q}}[\max(F(T,T)-K,0)] \\ P(t,T,K) &= \exp(-r_t\tau)E_t^{\mathbb{Q}}[\max(K-F(T,T),0)], \end{aligned} \tag{6}$$

where r_t is the risk-free rate at time t and $\tau = T - t$ is the time to maturity. Note that the value of a futures contract should be equal to the VIX on its expiration date: $F(T,T) = \mathrm{VIX}_T$. Because of this relation, the prices of call and put options can be determined by the characteristic function of v_t.

[7] Note that the ODEs are state-independent so we do not need to repeat running a numerical ODE solver on every date for a given parameter set. In other words, given a parameter set, we numerically solve the ODEs just once and save them in a storage space. The stored ODE solution can be repeatedly used to compute the characteristic function on each date throughout the sample period.

Proposition 3. *$C(t,T,K)$ and $P(t,T,K)$ are given by*

$$\begin{aligned} C(t,T,K) &= \exp(-r_t\tau)\left[g(v_t,u_t,w_t,t,T;\phi=-\mathrm{i})\Pi_1(t,T,K) - K\Pi_2(t,T,K)\right] \\ P(t,T,K) &= \exp(-r_t\tau)\left[K(1-\Pi_2(t,T,K)) - g(v_t,u_t,w_t,t,T;\phi=-\mathrm{i})(1-\Pi_1(t,T,K))\right], \end{aligned} \tag{7}$$

where the complementary distribution functions, $\Pi_1(t,T,K)$ and $\Pi_2(t,T,K)$, may be expressed as

$$\begin{aligned} \Pi_1(t,T,K) &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \Re\left[\frac{\exp(-\mathrm{i}\phi\log(K))h(v_t,u_t,w_t,t,T;\phi)}{\mathrm{i}\phi}\right]d\phi \\ \Pi_2(t,T,K) &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \Re\left[\frac{\exp(-\mathrm{i}\phi\log(K))g(v_t,u_t,w_t,t,T;\phi)}{\mathrm{i}\phi}\right]d\phi, \end{aligned} \tag{8}$$

where $h(v_t,u_t,w_t,t,T;\phi) = g(v_t,u_t,w_t,t,T;\phi-\mathrm{i})/g(v_t,u_t,w_t,t,T;-\mathrm{i})$ and $\Re$ indicates the real part of a complex number.[8]

Proof. See Appendix A. □

$\Pi_2(t,T,K)$ indicates the probability that a call option with a maturity of T and a strike price of K ends up in the money on its expiration date. Among different kinds of numerical integration, we choose to use the Gauss-Laguerre quadrature method with an order of 20 in our application.

2.4 Estimation

The option pricing literature often estimates the models using a quasi maximum likelihood estimation based on a variant of Kalman filtering (see, for example, Carr and Wu (2003, 2007), Trolle and Schwartz (2009a,b), and Mencía and Sentana (2012)). Christoffersen, Dorion, Jacobs, and Karoui (2014) show that unscented Kalman filtering is superior to extended Kalman filtering in the application of interest rate derivatives and comparable to particle filtering, which is computationally far more

[8] One may want to replace the term $g(v_t,u_t,w_t,t,T;\phi=-\mathrm{i})$ with the observed futures prices (see, for example, the footnote 10 of Mencía and Sentana (2012)). However, we do not take this approach for several reasons. First, we find that the observed VIX future prices have often departed from those implied by the put-call parity condition to a large extent, although this departure is not reported in this paper. Second, when the term $g(v_t,u_t,w_t,t,T;\phi=-\mathrm{i})$ is replaced with the observed prices, we face the numerical issue that the correlation parameter is estimated to be virtually one. Lastly, using the observed future prices in lieu of $g(v_t,u_t,w_t,t,T;-\mathrm{i})$ cannot guarantee the positivity of the option pricing formula in Equation (7).

expensive. Hence, in this paper, we apply a quasi maximum likelihood method via an unscented Kalman filter.[9] The Kalman filter requires us to recast the models into a state space form that comprises state equations and observation equations. First, it is necessary to introduce state equations for (u_t, w_t) for the physical $\mathbb{P}$ measure. To do so, we assume that u_t and w_t take the following physical dynamics:

$$\begin{aligned} du_t &= \kappa_u(\overline{u} - u_t)dt + \eta_u u_t dt + \sigma_u dB_{2t}^{\mathbb{P}} \\ dw_t &= \kappa_w(\overline{w} - w_t)dt + \eta_w w_t dt + \sigma_w \sqrt{w_t} dB_{3t}^{\mathbb{P}}, \end{aligned} \tag{9}$$

where $B_{2t}^{\mathbb{P}}$ and $B_{3t}^{\mathbb{P}}$ are independent standard Brownian motions under the physical $\mathbb{P}$ measure and $\eta_u u_t$ and $\eta_w w_t$ capture the risk premiums for the u_t and w_t processes, respectively. Note that we do not need the physical dynamics for v_t because our observation equations involve only the risk-neutral dynamics. For this reason, we are unable to pin down jump risk premiums.

By applying an Euler approximation to Equation (9), we are able to define discrete-time state equations:

$$\begin{aligned} u_{t+\Delta} &= u_t + \kappa_u(\overline{u} - u_t)\Delta + \eta_u u_t \Delta + \sigma_u \sqrt{\Delta}\varepsilon_{2,t+\Delta} \\ w_{t+\Delta} &= w_t + \kappa_w(\overline{w} - w_t)\Delta + \eta_w w_t \Delta + \sigma_w \sqrt{w_t \Delta}\varepsilon_{3,t+\Delta}, \end{aligned} \tag{10}$$

where Δ is the time interval and $\varepsilon_{2,t+\Delta}$ and $\varepsilon_{3,t+\Delta}$ are independent standard normal random variables.

With respect to observation equations, we assume that VIX futures and options are observed with measurement errors. Specifically, following Trolle and Schwartz (2009b), we assume that measurement errors in log future prices are constant and that measurement errors in options prices are proportional to their market vegas, denoted by $\widehat{\nu}(t, T, K)$. That is,

$$\begin{aligned} \log(\widehat{F}(t,T)) &= \log(F(t,T)) + \sigma_F \xi_{1,t} \\ \widehat{C}(t,T,K) &= C(t,T,K) + \widehat{\nu}(t,T,K)\sigma_E \xi_{2,t} \\ \widehat{P}(t,T,K) &= P(t,T,K) + \widehat{\nu}(t,T,K)\sigma_E \xi_{3,t}, \end{aligned} \tag{11}$$

where the hat symbol indicates the market price; σ_F and σ_E capture the sizes of measurement errors for futures and options, respectively; and $\xi_{1,t}$, $\xi_{2,t}$, and $\xi_{3,t}$ are in-

[9] We also implement an extended Kalman filter and find that the results are robust to the type of a Kalman filter.

dependent standard normal random variables. Details on unscented Kalman filtering are described in Appendix C.

2.5 Mencía and Sentana (2012) model

Recently, Mencía and Sentana (2012) studied the impact of stochastic volatility on the valuation of VIX derivatives. Although ours is not the first paper to address the importance of this subject, we take a different approach to modeling stochastic volatility than Mencía and Sentana (2012) do. That is, stochastic volatility is diffusion driven in our models, whereas it is jump driven in the Mencía and Sentana (2012) model. To compare the diffusion-driven model to the jump-driven model, we estimate one of the extended models of Mencía and Sentana (2012), which is hereafter referred to as the MS model. The MS model is specified under the risk-neutral measure as

$$\begin{aligned} dv_t &= \kappa_v(u_t - v_t)dt + \sqrt{z_t}dB_{1t}^{\mathbb{Q}} \\ du_t &= \kappa_u(\overline{u} - u_t)dt + \sigma_u dB_{2t}^{\mathbb{Q}} \\ dz_t &= -\lambda_z z_t dt + dq_t^{\mathbb{Q}}, \end{aligned} \tag{12}$$

where z_t is a volatility factor and $q_t^{\mathbb{Q}}$ is assumed to follow a compound Poisson process with intensity λ_z and an exponential jump distribution with a positive mean of δ_z. Note that the dynamics of z_t in Equation (12) belong to a class of the non-Gaussian Ornstein-Uhlenbeck process of Barndorff-Nielsen and Shephard (2001).

3 Data

Our application looks at the valuation of VIX futures and options, among other volatility derivatives. The markets for these derivatives have seen explosive growth in trading activity in recent years, as can be seen in Figure 2. The top panel of the figure shows that the number of VIX futures contracts traded increased dramatically from about 1 million in 2007 to about 24 million in 2012, and that most of the growth occurred after 2009, likely provoked by the recent financial crisis. The bottom panel of the figure shows that the dollar trading volume for the VIX options also increased substantially from around $3.3 billion in 2007 to around $16.1 billion in 2012. Note that VIX call options are more actively traded than VIX put options, which may be

associated with the fact that the former can be used as a means of hedging a stock market crash unlike the latter.

Our data set comprises the daily prices of VIX futures and options. The futures data come from Thomson Reuters, Datastream, while the options data are from OptionMetrics, Ivy DB. The VIX futures market started on March 26, 2004, and the VIX options market opened about two years later on February 24, 2006. As we decided to include only the dates for which data on both VIX futures and options are available, our sample starts on July 1, 2006, and ends on January 31, 2013, with a total of 1,657 trading days.[10] There are no early exercise premiums because VIX options are European.

In general, VIX futures are available for a wider range of maturities than VIX options. We delete VIX futures for which there are no paired VIX options with the same maturity. This leaves us a total of 9,330 futures contracts.

Option prices are taken from the bid–ask midpoint at each day's close of the options market (3:15 p.m). A filtering scheme is applied to eliminate inaccurate or illiquid options. Specifically, we delete VIX options for which the mid price is less than 0.1, the time to maturity is less than eight days, the Black-Scholes implied volatility is null, or the relative bid–ask spread (defined as the bid–ask spread divided by the mid price) is larger than 0.3. We also discard the VIX options that violate the lower bound constraints:

$$\begin{aligned} \widehat{C}(t,T,K) &\geq \max(0, \exp(-r_t\tau)(\widehat{F}(t,T) - K)) \\ \widehat{P}(t,T,K) &\geq \max(0, \exp(-r_t\tau)(K - \widehat{F}(t,T))), \end{aligned} \tag{13}$$

where the risk-free rates are obtained from OptionMetrics, Ivy DB. Option prices comprise intrinsic and time values, and only the time value portion is subject to model specifications. Moreover, OTM options tend to be more liquid than in-the-money ones. For this reason, we use only OTM call and put options in our empirical tests. Finally, a total of 102,525 options are included in our sample.

Table 2 shows summary statistics such as the number of observations and the average price for VIX futures and options. In this table, the data are divided into two categories depending on time to maturity τ: short-term contracts with $\tau \leq$

[10] Because the trading of VIX options was inactive in the very first few months, we exclude those months.

three months and long-term contracts with $\tau >$ three months. For options, we further break down the data into four categories according to their moneyness: deep OTM puts with $d < 0.85$, moderate OTM puts with $0.85 < d < 1.0$, moderate OTM calls with $1.0 < d < 1.3$, and deep OTM calls with $d > 1.3$, where moneyness is defined as $d = K/F(t,T)$. Note that call options have a larger number of observations and higher levels of Black-Scholes implied volatility than put options.

Figure 3 shows the time series of the VIX, the VVIX, and the option-implied skewness. The VIX and the VVIX are provided by the CBOE, while the one-month option-implied skewness is computed using a modified version of the model-free approach of Bakshi, Kapadia, and Madan (2003). The detailed procedure is provided in Appendix D. Two features are worth noting. First, the VIX and the VVIX share some pronounced spikes, such as the bankruptcy of Lehman Brothers, which is associated with asymmetric volatility. Second, the implied skewness takes a positive value most of the time, implying that a nonnormal stochastic mechanism is essential for describing the return distribution implied by VIX options.

4 Parameter estimates and preliminary analysis

This section starts by looking at two sets of parameter estimates. The first set is obtained using the entire data set from July 2006 through January 2013, whereas the second set is obtained using a subsample of four years from July 2006 through June 2010. The second set is used to test different model specifications for an out-of-sample period from July 2010 through January 2013. In addition, we will examine whether asymmetric volatility and jumps help to match the observed term structure of the implied volatility and skewness, on average, which will give us a hint as to the pricing performance results discussed in the section that follows.

4.1 Parameter estimates

Table 3 presents in-sample parameter estimates across various model specifications. The models with asymmetry in volatility yield very different volatility dynamics than the model without it. To see this, let us compare the SVV model with the AVV-type models (AVV, AVV-UJ, and AVV-AJ). First, we observe that the volatility process is

not very persistent in the SVV model ($\kappa_w = 5.506$), whereas it is more persistent in the AVV-type models ($0.847 \leq \kappa_w \leq 1.553$). Second, the SVV model ($\sigma_w = 7.271$) has far higher volatility of volatility than the AVV-type models ($1.976 \leq \sigma_w \leq 2.610$). Third, we define the long-run mean of volatility under the physical measure as $\overline{w}^{\mathbb{P}} = \frac{\kappa_w \overline{w}}{\kappa_w - \eta_w}$ and find that the SVV model ($\overline{w}^{\mathbb{P}} = 0.262$) has a lower level of the long-run mean than the AVV-type models ($0.553 \leq \overline{w}^{\mathbb{P}} \leq 0.677$). Overall, allowing for asymmetric volatility makes a large difference in the estimation of the volatility dynamics of the VIX, which in turn will have a large effect on the pricing of VIX derivatives.

The degree of asymmetry in volatility differs depending on whether upward and downward jumps are included. The asymmetry parameter is highest in the AVV model ($\rho = 0.870$) and lowest in the AVV-UJ model ($\rho = 0.422$). The AVV-AJ model ($\rho = 0.794$) falls in between. Because of this result, a jump process would have an impact on the pricing of VIX derivatives not only per se but also through different degrees of asymmetric volatility.

Interestingly, the variance risk premium parameter, η_w, takes a negative value for each of the models considered, implying that the market volatility of volatility is negatively priced. A similar discovery is also made by Song (2012) and Park (2015) in a non-parametric fashion. This negative pricing is not surprising because a market volatility of volatility asset can be used as a hedge against stock market downturns as it has a negative exposure (or beta) to the stock market index.

Now let us look at the frequencies and sizes of upward and downward jumps in the AVV-UJ and AVV-AJ models. The table shows that upward jumps take place about 2.7 times per annum with an average size of about 0.27, regardless of whether downward jumps are included. As expected, in the AVV-AJ model, upward jumps have a higher occurrence rate and a larger size than downward jumps; that is, $\lambda_+ > \lambda_-$ and $\delta_+ > |\delta_-|$.

Table 3 also contains two model comparison metrics: Akaike (AIC) and Schwarz (SIC) information criteria. These are based on some penalty scores, which depend on the number of free parameters, less log likelihood values. The lower the information criteria, the better the model. Both information criteria offer decisive rankings of the models: AVV-AJ > AVV-UJ > AVV > SVV > MS. In short, asymmetric volatility, upward jumps, and downward jumps are all critical modeling features, and the SVV model is preferred to the MS model.

Each panel of Figure 4 shows a time series of the filtered volatility states that are obtained using the in-sample parameter estimates. The filtered volatility states in the SVV model tend to be higher and more volatile than those in the AVV model. In addition, the filtered volatility states are substantially higher in the models with no jumps (SVV and AVV) than in the models with jumps (AVV-UJ and AVV-AJ), although all models have similar evolutions over time.

Table 4 presents out-of-sample parameter estimates across various model specifications. This table also indicates that the decision to allow for asymmetry in volatility has a large effect on the volatility dynamics and the decision to include jumps affects the extent of asymmetry in volatility.

4.2 Fitting the term structure of the implied volatility

Here, we investigate the ability of the models to fit the average term structure of the option-implied volatility. The option-implied volatility is computed by applying the model-free approach of Bakshi, Kapadia, and Madan (2003) to a cross section of VIX options, with little modification. The detailed procedure is provided in Appendix D.

The top panel of Figure 5 shows the scatter plot and kernel regression (solid line) of the implied volatility against months to maturity; the kernel regression is applied to capture the average term structure of the implied volatility. An interesting fact is that the term structure of the implied volatility is downward sloping, with the short end being nearly twice as high as the long end. This decreasing pattern is associated with the mean-reverting behavior of the VIX; the stronger the mean reversion of the VIX, the steeper the slope.

To see whether the models under consideration can capture the decreasing pattern in the implied volatility, we compute the average term structure of volatility implied by each model. The model-implied volatility, denoted by $\mathrm{IV}(t,T)$, can be computed by differentiating the cumulant generating function for each model:

$$\mathrm{IV}(t,T) = \sqrt{\frac{1}{\mathrm{i}^2}\frac{\partial^2 \Psi(t,T;\phi)/\partial\phi^2|_{\phi=0}}{(T-t)}}, \tag{14}$$

where the cumulant generating function is defined as $\Psi(t,T;\phi) = \log(g(v_t,u_t,w_t,t,T;\phi))$. Note that $\mathrm{IV}(t,T)$ is annualized. The model-implied volatility is computed on each

day and for each time to maturity throughout the sample period using the in-sample parameter estimates in Table 3. We then run a kernel regression of the model-implied volatility onto months to maturity for each model.

The bottom panel of Figure 5 shows that all of the four models—SVV (solid line), AVV (dashed line), AVV-UJ (dotted line), and AVV-AJ (dashed-dotted line)—do a good job of capturing the decreasing pattern in the observed implied volatility, although there are some minor discrepancies across the models. In particular, the SVV model tends to slightly overestimate the term structure of volatility, whereas the AVV-type models tend to slightly underestimate it. Overall, the models considered in this paper are all good at capturing the average term structure of the implied volatility, and the decision whether to allow for asymmetric volatility and jumps does not make a notable difference in this regard.

4.3 Fitting the term structure of the implied skewness

This subsection addresses whether asymmetric volatility and jumps can help explain the average term structure of the implied skewness. As for the option-implied volatility, the option-implied skewness is computed by applying the model-free approach of Bakshi, Kapadia, and Madan (2003) to a cross section of VIX options, with little modification.

The top panel of Figure 6 shows the scatter plot and kernel regression (solid line) of the implied skewness against months to maturity; the kernel regression is applied to capture the average term structure of the implied skewness. This figure reveals a surprising stylized fact: the average term structure of the implied skewness is nearly flat up to a one-year horizon.[11] Put simply, the short-term skewness is approximately as high as the long-term skewness.

To examine whether the models under consideration can produce such a flat pattern, we compute the model-implied skewness. Similar to the model-implied volatility, the model-implied skewness, denoted by $\mathrm{SKEW}(t,T)$, can be computed by differen-

[11] This finding echoes Carr and Wu (2003), who find that the stock index options market has the almost same slope of the implied volatility smile across different times to maturity. To imitate this pattern, they propose a finite moment log stable process.

tiating the cumulant generating function:

$$\text{SKEW}(t,T) = \frac{1}{\mathrm{i}^3} \frac{\partial^3 \Psi(t,T;\phi)/\partial \phi^3\big|_{\phi=0}}{(T-t)^{3/2}\text{IV}(t,T)^3}. \tag{15}$$

The bottom panel of Figure 6 shows that the AVV (dashed line) and AVV-UJ (dotted line) models both do a fairly good job of capturing the flat term structure of skewness. The AVV-UJ model tends to generate more of a positive skewness for short horizons than the AVV model, implying that upward jumps have a greater effect for short horizons than for long ones. In contrast, the AVV-AJ model (dashed-dotted line) does a poorer job of fitting the term structure of the implied skewness than the AVV and AVV-UJ models. Finally, the SVV model (solid line) has no ability to produce a positive skewness at all, (although it is capable of generating a positive excess kurtosis).

5 Pricing performance

Our performance analysis focuses on four main comparisons. First, we compare the diffusion-driven SVV model with the jump-driven MS model. Second, we compare the AVV model with the SVV model to test the importance of allowing for asymmetry in volatility. The third comparison is made between the AVV-UJ model and the AVV model in order to examine the effects of including upward jumps on top of asymmetric volatility. Lastly, we compare the AVV-AJ model with the AVV-UJ model to investigate whether the addition of downward jumps can make an incremental improvement once asymmetric volatility and upward jumps are included.

To facilitate our analysis, we compute a performance metric, a root mean squared error (RMSE), which is separately defined for futures and options as

$$\begin{aligned} \text{RMSE}_i^F &= \sqrt{\frac{1}{N^F}\sum_{n=1}^{N^F}\Big(\log(\widehat{F}(t_n,T_n)) - \log(F(t_n,T_n))\Big)^2} \\ \text{RMSE}_i^O &= \sqrt{\frac{1}{N^O}\sum_{n=1}^{N^O}\Big(\frac{\widehat{O}(t_n,T_n,K_n) - O(t_n,T_n,K_n)}{\widehat{\nu}(t_n,T_n,K_n)}\Big)^2}, \end{aligned} \tag{16}$$

where RMSE_i^F and RMSE_i^O denote the RMSEs implied by a model i for futures and

options, respectively; N^F and N^O indicate the total numbers of futures and options included in the sample, respectively; and $\widehat{O}_i$ and O_i denote the market-implied and model-implied option prices, respectively. Hereafter, RMSE_i is used to denote either RMSE_i^F or RMSE_i^O according to the context of analysis. Note that the performance metric is consistent with the assumption that we make about measurement errors in the observation equations.

To measure the extent to which one model is better or worse than another, we compute pricing differences between two models. Let $\Delta\text{RMSE}_{i|j}$ denote the pricing difference of a model i over a model j. $\Delta\text{RMSE}_{i|j}$ is defined as

$$\Delta\text{RMSE}_{i|j} = 100 \times \log(\text{RMSE}_i/\text{RMSE}_j), \tag{17}$$

where RMSE_i and RMSE_j denotes the RMSEs implied by models i and j, respectively. A negative (positive) value of $\Delta\text{RMSE}_{i|j}$ means that model i yields lower (higher) pricing errors than model j, implying that the pricing performance of the former is better (worse) than that of the latter by a percentage of that value.

Tables 5 and 6 show in-sample and out-of-sample pricing performance, respectively, across the different models. From left to right, the groups of columns show the pricing errors as measured by RMSE_i; the pricing improvements relative to the SVV model as measured by $\Delta\text{RMSE}_{i|\text{SVV}}$; the pricing improvements relative to the AVV model as measured by $\Delta\text{RMSE}_{i|\text{AVV}}$; and the pricing improvements relative to the AVV-UJ model as measured by $\Delta\text{RMSE}_{i|\text{AVV-UJ}}$, respectively. The main results for all futures and options are shown in the top two panels, and the options are broken down into calls and puts and by moneyness levels in the lower panels.

To assess the statistical significance of a pricing performance difference, we make pairwise model comparisons based on the Diebold and Mariano (2002) test. To do so, we compute a time series of RMSEs for each model, $\{\text{RMSE}_{i,t}\}_{t=1}^D$, where D is the total number of evaluation days and $\text{RMSE}_{i,t}$ is the date-t RMSE implied by a model i. Note that $\{\text{RMSE}_{i,t}\}_{t=1}^D$ is separately obtained for in-sample and out-of-sample tests.

Next, we obtain a series of the differences in RMSEs between every pair of the models, $d_{i,j,t} = \text{RMSE}_{i,t} - \text{RMSE}_{j,t}$, $t = 1, \ldots, D$, and define a test statistic between

models i and j as

$$z_{i,j} = \frac{\overline{d_{i,j,t}}}{\sigma(d_{i,j,t})}, \tag{18}$$

where $\overline{d_{i,j,t}}$ is a sample mean of the differences and $\sigma(d_{i,j,t})$ is a heteroskedasticity- and autocorrelation-consistent estimate of the sample standard deviation of the differences.[12] Diebold and Mariano (2002) show that the test statistic in Equation (18) follows a standard normal distribution under the null hypothesis that there is no statistically significant difference between models i and j.

Tables 7 and 8 display in-sample and out-of-sample test statistics, respectively. A negative (positive) statistic in a cell (i, j) indicates that model i outperforms (underperforms) model j. Panels A and B correspond to VIX futures and options, respectively. The symbols *, **, and *** represent statistical significance at the 10, 5, and 1 percent levels, respectively.

5.1 Diffusion-driven versus jump-driven volatility

This subsection focuses on investigating whether the SVV model can price VIX futures and options better than the MS model. With respect to futures, the pricing errors and pairwise model comparisons are shown in Panel A of Tables 5–8. The SVV model has an in-sample RMSE of 0.1842 and an out-of-sample RMSE of 0.2252, while the MS model has an in-sample RMSE of 0.1853 and an out-of-sample RMSE of 0.2332. The $\Delta\text{RMSE}_{i|\text{SVV}}$ metrics indicate that the futures pricing performance of the SVV model is better than that of the MS model by 1.18 percent in the in-sample test and by 7.02 percent in the out-of-sample test. The pricing difference is statistically significant at the 1 percent level only in the out-of-sample test, with a z statistic of ± 7.58 .

With respect to options, the pricing errors and pairwise model comparisons are shown in Panel B of Tables 5–8. The SVV model has an in-sample RMSE of 0.0939 and an out-of-sample RMSE of 0.1199, while the MS model has an in-sample RMSE of 0.1037 and an out-of-sample RMSE of 0.1475. The $\Delta\text{RMSE}_{i|\text{SVV}}$ metrics indicate that the options pricing performance of the SVV model is better than that of the MS model by 10.00 percent in the in-sample test and by 20.71 percent in the out-of-

[12] To control for autocorrelations, we use an optimal number of lags following Newey and West (1994).

sample test. The performance difference is statistically significant at the 1 percent level with a z statistic of ± 6.44 in the in-sample test and at the 1 percent level with a z statistic of ± 12.26 in the out-of-sample test.

To summarize, the SVV model outperforms the MS model except for the in-sample futures pricing. The superiority of the SVV model over the MS model implies that the volatility of the VIX is better described by a diffusive process rather than by a finite-activity jump process.[13] For this reason, the SVV model is preferred over the MS model as a baseline model upon which to add asymmetric volatility and jumps.

5.2 Effects of asymmetric volatility

This subsection focuses on discussing the effects of allowing for asymmetry in volatility on the pricing of futures and options by comparing the AVV model with the SVV model. With respect to futures, the AVV model has an in-sample RMSE of 0.1804 and an out-of-sample RMSE of 0.2201. Based on the $\Delta\text{RMSE}_{i|\text{SVV}}$ metrics, the futures pricing performance of the AVV model is better than that of the SVV model by 4.22 percent in the in-sample test and by 4.58 percent in the out-of-sample test. The pricing difference is statistically significant at the 5 percent level with a z statistic of ± 2.52 in the in-sample test and at the 1 percent level with a z statistic of ± 8.17 in the out-of-sample test. Furthermore, the futures pricing improvements vary by time to maturity. Specifically, the AVV model generates better futures pricing than the SVV model for short-term horizons both in-sample and out-of-sample but not for long-term horizons.

With respect to options, the AVV model has an in-sample RMSE of 0.0885 and an out-of-sample RMSE of 0.1154. The $\Delta\text{RMSE}_{i|\text{SVV}}$ metrics suggest that the options pricing performance of the AVV model is better than that of the SVV model by 5.87 percent in the in-sample test and by 3.87 percent in the out-of-sample test. The performance difference is statistically significant at the 1 percent level with a z statistic of ± 7.08 in the in-sample test and at the 1 percent level with a z statistic of ± 6.49 in the out-of-sample test. Furthermore, we find that allowing for asymmetry in volatility reduces option pricing errors for all horizons in both in-sample and out-

[13] This paper considers only two specifications of stochastic volatility of the VIX: the MS and SVV models. A further analysis of other specifications, such as infinite-activity jumps or a combination of diffusive and jump processes, would be an interesting topic for future research.

of-sample tests.

Taking our analysis one step further, we break down option pricing errors into puts and calls. Panels C and D of Table 5 show the in-sample RMSEs for OTM puts and calls, respectively. Compared to the SVV model, the AVV model substantially improves the pricing of put options by 14.60 percent but improves the pricing of call options only by 1.02 percent. Similarly, Panels C and D of Table 6 show the out-of-sample RMSEs for OTM puts and calls, respectively. Again, compared with the SVV model, the AVV model makes a large improvement in the pricing of put options by 11.45 percent, but only a very small improvement in the pricing of call options by 1.86 percent.

To sum up, we find decisive evidence that asymmetry in volatility has an enormous and statistically significant effect on the valuation of VIX futures and options. The AVV model is strongly preferred to the SVV model in both in-sample and out-of-sample tests. Especially, asymmetry in volatility plays a crucial role in fitting the prices of puts and short-term futures. Given the fact that OTM put prices are usually found to be cheaper than OTM call prices with an equivalent moneyness, our result implies that the underpricing of VIX puts is more likely to be driven by asymmetric volatility rather than upward jumps.

5.3 Effects of upward jumps

We now investigate the effects of including upward jumps on top of asymmetric volatility on the pricing of futures and options by comparing the AVV-UJ model with the AVV model. With respect to futures, the AVV-UJ model has an in-sample RMSE of 0.1799 and an out-of-sample RMSE of 0.2163. The $\Delta\text{RMSE}_{i|\text{AVV}}$ metrics indicate that the futures pricing performance of the AVV-UJ model is better than that of the AVV model by 0.53 percent in the in-sample test and by 3.51 percent in the out-of-sample test. Hence, the statistical significance is obtained only for the out-of-sample test at the 1 percent level with a z statistic of ± 6.50.

Furthermore, the futures pricing improvement depends on time to maturity in both in-sample and out-of-sample tests. Incorporating upward jumps improves the pricing of long-term futures but worsens the pricing of short-term futures. This finding can be contrasted with the previous finding that allowing for asymmetry in

volatility helps to improve the pricing of short-term contracts only. Taken together, asymmetric volatility and upward jumps have opposing yet complementary influences on futures pricing performance.

With respect to options, the AVV-UJ model has an in-sample RMSE of 0.0817 and an out-of-sample RMSE of 0.1094. The $\Delta\text{RMSE}_{i|\text{AVV}}$ metrics suggest that the options pricing performance of the AVV-UJ model is better than that of the AVV model by 7.96 percent in the in-sample test and by 5.33 percent in the out-of-sample test. The performance difference is statistically significant at the 1 percent level with a z statistic of ± 8.16 in the in-sample test and at the 1 percent level with a z statistic of ± 11.71 in the out-of-sample test. Furthermore, we find that including upward jumps is effective regardless of option maturities in both in-sample and out-of-sample tests, although the improvement is greater for long-term options than for short-term ones.

We now break down option pricing errors into puts and calls. Panels C and D of Tables 5 and 6 show the in-sample and out-of-sample RMSEs for OTM puts and calls, respectively. The AVV-UJ model greatly improves the pricing of cal in both in-sample and out-of-sample tests, but there is no significant pricing difference between the two models with respect to the pricing of put options.

Based on the empirical results above, we conclude that upward jumps have a sizable and statistically significant impact on the valuation of VIX futures and options. The AVV-UJ model is strongly preferred to the AVV model in both in-sample and out-of-sample tests, except for the in-sample futures pricing. Especially, upward jumps play a pivotal role in pricing calls and long-term futures. Given the fact that OTM call prices are usually more expensive than OTM put prices with an equivalent moneyness, our result implies that high levels of call prices are more likely to be driven by upward jumps rather than by asymmetric volatility.

5.4 Effects of downward jumps

Lastly, we evaluate the incremental effect of downward jumps on the pricing of futures and option by comparing the AVV-AJ model with the AVV-UJ model. Based on the $\Delta\text{RMSE}_{i|\text{AVV-UJ}}$ metrics, including downward jumps makes little improvement in the pricing of futures and options to the extent that asymmetric volatility and

upward jumps are already accounted for. The statistical significance is obtained only for in-sample futures pricing but the size of improvement is quite small with $\Delta\text{RMSE}_{i|\text{AVV-UJ}} = -0.86$.

To see if the inefficacy of downward jumps relies on our particular choice of jump specifications, we also consider an alternative specification of upward and downward jumps. In particular, we estimate the AVV model with double exponential jumps as in Kou (2002) and Kou and Wang (2004), where upward and downward jumps are driven by a single compound Poisson process (see Appendix E for details on the model specification). Once again, we fail to find evidence that the AVV model with double exponential jumps is better at pricing the VIX derivatives than the AVV-UJ model, although the results are not reported in this paper.

6 Conclusion

Given the growing interest in volatility trading, this paper suggests a class of affine jump-diffusion models for the pricing of VIX derivatives. The models have two innovative features. First, to assess the effects of asymmetric volatility on the pricing of VIX derivatives, we allow for a positive correlation between changes in the VIX and in its stochastic volatility. Second, we separate out the roles of upward and downward jumps to better consider the reality that investors react differently to good and bad shocks, which may arrive independently with different rates and sizes.

Our empirical analysis shows decisive evidence for the benefits of including both asymmetric volatility and upward jumps in pricing models of VIX futures and options. In particular, we find that asymmetric volatility is essential for pricing put options and short-term futures, whereas upward jumps are crucial for pricing call options and long-term futures. That is, low levels of put prices are associated with asymmetric volatility, whereas high levels of call prices are likely to be driven by the possibility of upward jumps. In sum, asymmetric volatility and upward jumps both are required to price VIX derivatives, having opposing yet complementary effects on them.

However, we fail to find consistent evidence for the benefit of including downward jumps in the pricing of the VIX futures and options. Given that all the jump specifications that we consider are of finite activity, we do not interpret our results as implying that downward jumps are of no use in explaining the VIX derivative

prices. Indeed, authors such as Todorov and Tauchen (2011) and Wu (2011) show that volatility jumps may be better described by an infinite-activity jump process than a finite-activity one. Therefore, it would be interesting to extend our research to models with infinite-activity jump processes (on top of asymmetric volatility).

Lastly, we report some interesting features of the term structures of the implied volatility and skewness. Specifically, the term structure of the implied volatility has a downward slope, whereas the term structure of the implied skewness is nearly flat. Although all of the models considered do a nice job of capturing the decreasing pattern in the implied volatility, only the AVV and AVV-UJ models successfully match the flat pattern in the implied skewness, consistent with the result that asymmetric volatility and upward jumps are both crucial for pricing the VIX options.

References

Aït-Sahalia, Y., D. Amengual, and E. Manresa (2015): "Market-based estimation of stochastic volatility models," *Journal of Econometrics*, 187, 418–435.

Aït-Sahalia, Y., and R. Kimmel (2007): "Maximum likelihood estimation of stochastic volatility models," *Journal of Financial Economics*, 83, 413–452.

Alexander, C., and D. Korovilas (2012): "Understanding ETNs on VIX futures," Working paper, University of Reading.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and C. Vega (2007): "Real-time price discovery in global stock, bond and foreign exchange markets," *Journal of International Economics*, 73(2), 251–277.

Bakshi, G., C. Cao, and Z. Chen (1997): "Empirical performance of alternative option pricing models," *Journal of Finance*, 52, 527–566.

Bakshi, G., P. Carr, and L. Wu (2008): "Stochastic risk premiums, stochastic skewness in currency options, and stochastic discount factors in international economies," *Journal of Financial Economics*, 87(1), 132–156.

Bakshi, G., N. Kapadia, and D. Madan (2003): "Stock return characteristics, skew laws, and differential pricing of individual equity options," *Review of Financial Studies*, 16, 101–143.

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

Bates, D. (2000): "Post-'87 crash fears in S&P500 futures options," *Journal of Econometrics*, 94, 181–238.

Berndt, E., B. Hall, R. Hall, and J. Hausman (1974): "Estimation and inference in nonlinear structural models," *Annals of Economic and Social Measurement*, 3, 653–665.

Bikbov, R., and M. Chernov (2009): "Unspanned stochastic volatility in affine models: Evidence from Eurodollar futures and options," *Management Science*, 55(8), 1292–1305.

Broadie, M., M. Chernov, and M. Johannes (2007): "Model specification and risk premia: Evidence from futures options," *Journal of Finance*, 62, 1453–1490.

Carr, P., and L. Wu (2003): "The finite moment log stable process and option pricing," *Journal of Finance*, 58, 753–777.

——— (2007): "Stochastic skew in currency options," *Journal of Financial Economics*, 86, 213–247.

——— (2009): "Variance risk premiums," *Review of Financial Studies*, 22, 1311–1341.

Christoffersen, P., C. Dorion, K. Jacobs, and L. Karoui (2014): "Nonlinear Kalman filtering in affine term structure models," *Management Science*, 60, 2248–2268.

Christoffersen, P., S. Heston, and K. Jacobs (2009): "The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well," *Management Science*, 55, 1914–1932.

Collin-Dufresne, P., and R. S. Goldstein (2002): "Do bonds span the fixed income markets? Theory and evidence for unspanned stochastic volatility," *Journal of Finance*, 57(4), 1685–1730.

Dai, Q., and K. J. Singleton (2000): "Specification analysis of affine term structure models," *Journal of Finance*, 55(5), 1943–1978.

Detemple, J., and C. Osakwe (2000): "The valuation of volatility options," *European Finance Review*, 4(1), 21–50.

Diebold, F. X., and R. S. Mariano (2002): "Comparing predictive accuracy," *Journal of Business and Economic Statistics*, 20(1), 134–144.

Duffie, D., J. Pan, and K. Singleton (2000): "Transform analysis and asset pricing for affine jump-diffusions," *Econometrica*, 68, 1343–1376.

Durham, G. B. (2013): "Risk-neutral modelling with affine and nonaffine models," *Journal of Financial Econometrics*, 11, 650–681.

Eraker, B. (2005): "Do stock prices and volatility jump? Reconciling evidence from spot and option prices," *Journal of Finance*, 59(3), 1367–1404.

Gallant, A. R., C. Hsu, and G. Tauchen (1999): "Using daily range data to calibrate volatility diffusions and extract the forward integrated variance," *Review of Economics and Statistics*, 81, 617–631.

Goard, J., and M. Mazur (2013): "Stochastic volatility models and the pricing of VIX options," *Mathematical Finance*, 23, 439–458.

Grünbichler, A., and F. A. Longstaff (1996): "Valuing futures and options on volatility," *Journal of Banking and Finance*, 20(6), 985–1001.

Heston, S. L. (1993): "A closed-form solution for options with stochastic volatility with applications to bond and currency options," *Review of Financial Studies*, 6, 327–343.

Huang, S., and J. Yu (2007): "On stiffness in affine asset pricing models," *Journal of Computational Finance*, 10(3), 99–123.

Jiang, G., and Y. Tian (2007): "Extracting model-free volatility from option prices: An examination of the VIX index," *Journal of Derivatives*, 14, 35–60.

Jones, C. (2003): "The dynamics of stochastic volatility: Evidence from underlying and options markets," *Journal of Econometrics*, 116, 181–224.

Julier, S. J. (2002): "The scaled unscented transformation," in *American Control Conference. Proceedings of the 2002*, vol. 6, pp. 4555–4559.

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

Kou, S. G., and H. Wang (2004): "Option pricing under a double exponential jump diffusion model," *Management Science*, 50(9), 1178–1192.

Li, H., and F. Zhao (2006): "Unspanned stochastic volatility: Evidence from hedging interest rate derivatives," *Journal of Finance*, 61(1), 341–378.

Mencía, J., and E. Sentana (2012): "Valuation of VIX derivatives," *Journal of Financial Economics*, 108, 367–391.

Newey, W. K., and K. D. West (1994): "Automatic lag selection in covariance matrix estimation," *Review of Economic Studies*, 61(4), 631–653.

Park, Y.-H. (2015): "Volatility of volatility and tail risk hedging returns," Forthcoming in Journal of Financial Markets, Federal Reserve Board.

Song, Z. (2012): "Expected VIX option returns," Working paper, Federal Reserve Board.

Todorov, V., and G. Tauchen (2011): "Volatility jumps," *Journal of Business and Economic Statistics*, 29(3), 356–371.

Todorov, V., G. Tauchen, and I. Grynkiv (2014): "Volatility activity: Specification and estimation," *Journal of Econometrics*, 178, 180–193.

Trolle, A. B., and E. S. Schwartz (2009a): "A general stochastic volatility model for the pricing of interest rate derivatives," *Review of Financial Studies*, 22(5), 2007–2057.

——— (2009b): "Unspanned stochastic volatility and the pricing of commodity derivatives," *Review of Financial Studies*, 22(11), 4423–4461.

Whaley, R. E. (2013): "Trading volatility: At what cost," *Journal of Portfolio Management*, 40(1), 95–108.

Wu, L. (2011): "Variance dynamics: Joint evidence from options and high-frequency returns," *Journal of Econometrics*, 160, 280–287.

A Derivation of the VIX derivative prices

We start by deriving the characteristic function, $g(v_t, u_t, w_t, t, T; \phi) = E_t^{\mathbb{Q}}\left[\exp(\mathrm{i}\phi v_T)\right]$, as in Heston (1993). Given the model in Equation (1), Ito's lemma leads to

$$\begin{aligned} dg = &\Big[g_t + g_v(\kappa_v(u-v) - \lambda_+\delta_+ - \lambda_-\delta_-) + g_u\kappa_u(\overline{u}-u) + g_w\kappa_w(\overline{w}-w)\Big]dt \\ &+ \Big[\tfrac{1}{2}g_{vv}w + \tfrac{1}{2}g_{uu}\sigma_u^2 + \tfrac{1}{2}g_{ww}\sigma_w^2 w + g_{vw}\rho\sigma_w w\Big]dt + g_v\sqrt{w}dB_{1t}^{\mathbb{Q}} + g_u\sigma_u dB_{2t}^{\mathbb{Q}} + g_w\sigma_w\sqrt{w}dB_{3t}^{\mathbb{Q}} \\ &+ [g(v+J_1^{\mathbb{Q}}, u, w, t, T; \phi) - g(v, u, w, t, T; \phi)]dN_{1t}^{\mathbb{Q}} \\ &+ [g(v+J_2^{\mathbb{Q}}, u, w, t, T; \phi) - g(v, u, w, t, T; \phi)]dN_{2t}^{\mathbb{Q}}. \end{aligned} \tag{A.1}$$

If a Markov process is affine, its cumulant generating function has affine dependence on the state vector:

$$g(v_t, u_t, w_t, t, T; \phi) = \exp\left[\alpha(s) + \beta_v(s)v_t + \beta_u(s)u_t + \beta_w(s)w_t\right]. \tag{A.2}$$

Substituting Equation (A.2) into Equation (A.1), we obtain

$$\begin{aligned} \frac{E_t[dg/g]}{dt} = &-[\dot{\alpha}(s) + \dot{\beta}_v(s)v + \dot{\beta}_u(s)u + \dot{\beta}_w(s)w] + [\kappa_v(u-v) - \lambda_+\delta_+ - \lambda_-\delta_-]\beta_v(s) \\ &+ \kappa_u(\overline{u}-u)\beta_u(s) + \kappa_w(\overline{w}-w)\beta_w(s) \\ &+ \tfrac{1}{2}w\beta_v(s)^2 + \tfrac{1}{2}\sigma_u^2\beta_u(s)^2 + \tfrac{1}{2}\sigma_w^2 w\beta_w(s)^2 + \rho\sigma_w w\beta_v(s)\beta_w(s) \\ &+ \lambda_+\left(E[\exp(\beta_v(s)J_1^{\mathbb{Q}})] - 1\right) + \lambda_-\left(E[\exp(\beta_v(s)J_2^{\mathbb{Q}})] - 1\right), \end{aligned} \tag{A.3}$$

where

$$\begin{aligned} E[\exp(\beta_v(s)J_1^{\mathbb{Q}})] &= \tfrac{1}{1-\delta_+\beta_v(s)} \\ E[\exp(\beta_v(s)J_2^{\mathbb{Q}})] &= \tfrac{1}{1-\delta_-\beta_v(s)}. \end{aligned} \tag{A.4}$$

Note that g is a martingale by the law of iterated expectations. Rearranging Equation (A.3) as an affine form of (v, u, w) and using the fact that $E_t[dg] = 0$, we

obtain the system of the following ODEs:

$$\begin{aligned}
\dot{\alpha}(s) &= \kappa_u \overline{u}\beta_u(s) + \kappa_w \overline{w}\beta_w(s) + \frac{1}{2}\sigma_u^2 \beta_u(s)^2 \\
&\quad - [\lambda_+ \delta_+ + \lambda_- \delta_-]\beta_v(s) + \lambda_+ \left[\frac{1}{1-\delta_+\beta_v(s)} - 1\right] + \lambda_- \left[\frac{1}{1-\delta_-\beta_v(s)} - 1\right] \\
\dot{\beta}_v(s) &= -\kappa_v \beta_v(s) \\
\dot{\beta}_u(s) &= \kappa_v \beta_v(s) - \kappa_u \beta_u(s) \\
\dot{\beta_w}(s) &= -\kappa_w \beta_w(s) + \frac{1}{2}\beta_v(s)^2 + \frac{1}{2}\sigma_w^2 \beta_w(s)^2 + \rho\sigma_w \beta_v(s)\beta_w(s),
\end{aligned} \tag{A.5}$$

where the boundary conditions are given as $\alpha(0) = 0$, $\beta_v(0) = \mathrm{i}\phi$, $\beta_u(0) = 0$, and $\beta_w(0) = 0$.

Let $F(t,T)$ denote the time-t price of a futures contract with a maturity of T. $F(t,T)$ is equal to the characteristic function evaluated at $\phi = -\mathrm{i}$:

$$\begin{aligned}
F(t,T) &= E_t^{\mathbb{Q}}[\mathrm{VIX}_T] \\
&= E_t^{\mathbb{Q}}[\exp(v_T)] \\
&= g(v_t, u_t, w_t, t, T; \phi = -\mathrm{i}).
\end{aligned} \tag{A.6}$$

To derive the pricing formula for $C(t,T,K)$, let us introduce a change of measure from $\mathbb{Q}$ to $\tilde{\mathbb{Q}}$ by a Radon-Nikodym derivative $\frac{d\tilde{\mathbb{Q}}}{d\mathbb{Q}} = \frac{\exp(v_T)}{E_t^{\mathbb{Q}}[\exp(v_T)]}$. Let us denote the characteristic function of v_T under the $\tilde{\mathbb{Q}}$ measure by $h(v_t, u_t, w_t, t, T; \phi)$. $h(v_t, u_t, w_t, t, T; \phi)$ is given by

$$\begin{aligned}
h(v_t, u_t, w_t, t, T; \phi) &= E_t^{\tilde{\mathbb{Q}}}[\exp(\mathrm{i}\phi v_T)] \\
&= E_t^{\mathbb{Q}}\left[\frac{\exp(v_T)}{E_t^{\mathbb{Q}}[\exp(v_T)]}\exp(\mathrm{i}\phi v_T)\right] \\
&= E_t^{\mathbb{Q}}\left[\frac{\exp(\mathrm{i}(\phi - \mathrm{i})v_T)}{E_t^{\mathbb{Q}}[\exp(v_T)]}\right] \\
&= \frac{g(v_t, u_t, w_t, t, T; \phi - \mathrm{i})}{g(v_t, u_t, w_t, t, T; -\mathrm{i})}.
\end{aligned} \tag{A.7}$$

$C(t,T,K)$ is calculated using $g(v_t,u_t,w_t,t,T;\phi)$ and $h(v_t,u_t,w_t,t,T;\phi)$:

$$\begin{aligned} C(t,T,K) &= \exp(-r_t\tau)E_t^{\mathbb{Q}}[\max(\text{VIX}_T - K, 0)] \\ &= \exp(-r_t\tau)E_t^{\mathbb{Q}}[\exp(v_T)\mathbb{1}_{(v_T\geq\log(K))} - K\mathbb{1}_{(v_T\geq\log(K))}] \\ &= \exp(-r_t\tau)\left[E_t^{\mathbb{Q}}[\exp(v_T)]E_t^{\tilde{\mathbb{Q}}}[\mathbb{1}_{(v_T\geq\log(K))}] - KE_t^{\mathbb{Q}}[\mathbb{1}_{(v_T\geq\log(K))}]\right] \\ &= \exp(-r_t\tau)\left[g(v_t,u_t,w_t,t,T;\phi=-\mathrm{i})\Pi_1(t,T,K) - K\Pi_2(t,T,K)\right], \end{aligned} \tag{A.8}$$

where $\Pi_1(t,T,K)$ and $\Pi_2(t,T,K)$ are given by the Levy inversion formula (see Appendix B):

$$\begin{aligned} \Pi_1(t,T,K) &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \Re\left[\frac{\exp(-\mathrm{i}\phi\log(K))h(v_t,u_t,w_t,t,T;\phi)}{\mathrm{i}\phi}\right]d\phi \\ \Pi_2(t,T,K) &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \Re\left[\frac{\exp(-\mathrm{i}\phi\log(K))g(v_t,u_t,w_t,t,T;\phi)}{\mathrm{i}\phi}\right]d\phi, \end{aligned} \tag{A.9}$$

where $\Re$ indicates the real part of a complex number. A similar procedure can be applied to derive the price of a put option.

B Levy inversion formula

Let $F(x)$ be the distribution function of a random variable. The Levy inversion formula involves computing $F(x)$ given the characteristic function, $g(\phi) = \int_{-\infty}^{\infty}\exp(\mathrm{i}\phi x)dF(x)$. To begin with, let us bring the following mathematical results:

$$\begin{aligned} \sin(x) &= \frac{\exp(\mathrm{i}x) - \exp(-\mathrm{i}x)}{2\mathrm{i}} \\ g(-\phi) &= \overline{g(\phi)} \\ \text{sgn}(x) &= \frac{2}{\pi}\int_0^\infty \frac{\sin(ux)}{u}du, \end{aligned} \tag{B.1}$$

where $\mathrm{sgn}(x) = -1$ if $x < 0$; 0 if $x = 0$; and 1 if $x > 0$. Let us derive the following relation:

$$\begin{aligned}
1 - 2F(x) &= -F(x) + (1 - F(x)) \\
&= -\int_{-\infty}^{x} dF(u) + \int_{x}^{\infty} dF(u) \\
&= \int_{-\infty}^{\infty} \mathrm{sgn}(u - x) dF(u) \\
&= \frac{2}{\pi} \int_{-\infty}^{\infty} \int_{0}^{\infty} \frac{\sin((u-x)\phi)}{\phi} d\phi dF(u) \\
&= -\frac{\mathrm{i}}{\pi} \int_{0}^{\infty} \int_{-\infty}^{\infty} \frac{\exp(\mathrm{i}(u-x)\phi) - \exp(-\mathrm{i}(u-x)\phi)}{\phi} dF(u) d\phi \\
&= -\frac{\mathrm{i}}{\pi} \int_{0}^{\infty} \frac{\exp(-\mathrm{i}x\phi) g(\phi) - \exp(\mathrm{i}x\phi) g(-\phi)}{\phi} d\phi \\
&= -\frac{\mathrm{i}}{\pi} \int_{0}^{\infty} \frac{\exp(-\mathrm{i}x\phi) g(\phi) - \overline{\exp(-\mathrm{i}x\phi) g(\phi)}}{\phi} d\phi \\
&= \frac{2}{\pi} \int_{0}^{\infty} \Im \left[\frac{\exp(-\mathrm{i}x\phi) g(\phi)}{\phi} \right] d\phi \\
&= \frac{2}{\pi} \int_{0}^{\infty} \Re \left[\frac{\exp(-\mathrm{i}x\phi) g(\phi)}{\mathrm{i}\phi} \right] d\phi,
\end{aligned} \tag{B.2}$$

where $\Re$ and $\Im$ denote the real and imaginary parts of a complex number, respectively. The distribution function is thus given by

$$F(x) = \frac{1}{2} - \frac{1}{\pi} \int_{0}^{\infty} \Re \left[\frac{\exp(-\mathrm{i}x\phi) g(\phi)}{\mathrm{i}\phi} \right] d\phi. \tag{B.3}$$

The option pricing formula calls for the complementary distribution function, $\Pi(x)$:

$$\begin{aligned}
\Pi(x) &= 1 - F(x) \\
&= \frac{1}{2} + \frac{1}{\pi} \int_{0}^{\infty} \Re \left[\frac{\exp(-\mathrm{i}x\phi) g(\phi)}{\mathrm{i}\phi} \right] d\phi.
\end{aligned} \tag{B.4}$$

C Unscented Kalman filtering

Let $X_t = (v_t, u_t)$ be the latent state vector. Suppose that the latent state vector is of the form:

$$X_{t+1} = \Phi_0 + \Phi_x X_t + \epsilon_{t+1}, \tag{C.1}$$

where $\text{var}_t[\epsilon_{t+1}] = Q(X_t)$. Kalman filtering comprises two stages: the prediction and the update stages. The prediction stage involves computing one-step-ahead predictive mean and variance of X_t given $\hat{X}_t = E_t[X_t]$ and $P_t = \text{var}_t[X_t]$:

$$\begin{aligned} \hat{X}_{t+1|t} &= \Phi_0 + \Phi_x \hat{X}_t \\ P_{t+1|t} &= \Phi_x P_t \Phi_x^{'} + Q(X_t) \end{aligned} \tag{C.2}$$

The update stage starts by defining a set of sigma points that match the predictive mean and variance of the latent vector. To be specific, a set of $2L+1$ sigma points and weights are selected based on the scaled unscented transformation of Julier (2002), where L is the dimension of the latent vector:

$$\begin{aligned} \mathscr{X}_0 &= \hat{X}_{t+1|t} \\ \mathscr{X}_i &= \hat{X}_{t+1|t} + \left(\sqrt{(L+\xi)P_{t+1|t}}\right)_i, \quad i = 1, \ldots, L \\ \mathscr{X}_i &= \hat{X}_{t+1|t} - \left(\sqrt{(L+\xi)P_{t+1|t}}\right)_{i-L}, \quad i = L+1, \ldots, 2L, \end{aligned} \tag{C.3}$$

whose weights are given by

$$\begin{aligned} W_0^m &= \frac{\xi}{L+\xi} \\ W_0^c &= \frac{\xi}{L+\xi} + (1 - \alpha^2 + \beta) \\ W_i^m &= \frac{1}{2(L+\xi)}, \quad i = 1, \ldots, 2L \\ W_i^c &= \frac{1}{2(L+\xi)}, \quad i = 1, \ldots, 2L, \end{aligned} \tag{C.4}$$

where $\xi = \alpha^2(L+\kappa) - L$ and $\left(\sqrt{(L+\xi)P_{t+1|t}}\right)_i$ is the ith column of the matrix square root. In our applications, we set $\alpha = 0.01$, $\beta = 2$, and $\kappa = 0$.

Suppose that the observation equations are of the form:

$$Y_t = H(X_t) + \Omega, \tag{C.5}$$

where $H(X_t)$ represents the pricing formula for futures and options and Ω denotes the measurement errors. The predictive mean and variance of Y_t are computed using

the sigma points and weights:

$$\begin{aligned}\hat{Y}_{t+1|t} &= \sum_{i=0}^{2L} W_i^m H(\mathscr{X}_i) \\ P^y_{t+1|t} &= \sum_{i=0}^{2L} W_i^c (H(\mathscr{X}_i) - \hat{Y}_{t+1|t})(H(\mathscr{X}_i) - \hat{Y}_{t+1|t})' + \Omega,\end{aligned} \tag{C.6}$$

where $\hat{Y}_{t+1|t} = E_t[Y_{t+1}]$ and $P^y_{t+1|t} = \text{var}_t[Y_{t+1}]$.

Finally, we are able to update the conditional mean and variance of X_t:

$$\begin{aligned}\hat{X}_{t+1} &= \hat{X}_{t+1|t} + K_{t+1}(Y_{t+1} - \hat{Y}_{t+1|t}) \\ P_{t+1} &= P_{t+1|t} - K_{t+1} P^y_{t+1|t} K'_{t+1},\end{aligned} \tag{C.7}$$

where

$$K_{t+1} = \sum_{i=0}^{2L} W_i^c (\mathscr{X}_i - \hat{X}_{t+1|t})(H(\mathscr{X}_i) - \hat{Y}_{t+1|t})'(P^y_{t+1|t})^{-1}. \tag{C.8}$$

The log likelihood function is then given by

$$\log(L) = -\frac{1}{2}\sum_{t=1}^{T}\left[\log(2\pi)N_t + \log\left(\left|\det(P^y_{t+1|t})\right|\right) + (Y_{t+1} - \hat{Y}_{t+1|t})'(P^y_{t+1|t})^{-1}(Y_{t+1} - \hat{Y}_{t+1|t})\right], \tag{C.9}$$

where N_t is the dimension of the observation equations at time t. The log likelihood function is optimized via the Berndt, Hall, Hall, and Hausman (1974) algorithm. The standard errors are computed by inverting the Hessian matrix at the optimal solution.

D Computation of the option-implied volatility and skewness

The model-free, option-implied moments are obtained with the approach of Bakshi, Kapadia, and Madan (2003) with light modification. The original formula is based on the no-arbitrage relation that the underlying asset is tradable and so grows at a risk-free rate under the risk-neutral measure. That is, $E_t^{\mathbb{Q}}(S_T) = S_t \exp(r_t(T-t))$, where S_t is the price of an underlying asset at time t. However, this relation no longer holds in the VIX derivatives market because the underlying VIX index is not tradable. Instead, we use the fact that a VIX futures price reflects the risk-neutral

expectation of the VIX, $F(t,T) = E_t^{\mathbb{Q}}(\text{VIX}_T)$, where $F(t,T)$ is the time-t price of a VIX future with a maturity of T. The computation proceeds as follows.

Let $H_1(t,T)$, $H_2(t,T)$, and $H_3(t,T)$ denote the prices of three hypothetical securities that pay quadratic, cubic, and quartic payoffs, respectively. These prices are given by

$$\begin{aligned} H_1(t,T) &= \int_{\text{VIX}_t}^{\infty} \frac{2\left(1-\ln\left(\frac{K}{\text{VIX}_t}\right)\right)}{K^2}\widehat{C}(t,T,K)dK \\ &\quad + \int_0^{\text{VIX}_t} \frac{2\left(1+\ln\left(\frac{\text{VIX}_t}{K}\right)\right)}{K^2}\widehat{P}(t,T,K)dK, \end{aligned} \tag{D.1}$$

$$\begin{aligned} H_2(t,T) &= \int_{\text{VIX}_t}^{\infty} \frac{6\ln\left(\frac{K}{\text{VIX}_t}\right) - 3\left(\ln\left(\frac{K}{\text{VIX}_t}\right)\right)^2}{K^2}\widehat{C}(t,T,K)dK \\ &\quad - \int_0^{\text{VIX}_t} \frac{6\ln\left(\frac{\text{VIX}_t}{K}\right) + 3\left(\ln\left(\frac{\text{VIX}_t}{K}\right)\right)^2}{K^2}\widehat{P}(t,T,K)dK, \end{aligned} \tag{D.2}$$

and

$$\begin{aligned} H_3(t,T) &= \int_{\text{VIX}_t}^{\infty} \frac{12\left(\ln\left(\frac{K}{\text{VIX}_t}\right)\right)^2 - 4\left(\ln\left(\frac{K}{\text{VIX}_t}\right)\right)^3}{K^2}\widehat{C}(t,T,K)dK \\ &\quad + \int_0^{\text{VIX}_t} \frac{12\left(\ln\left(\frac{\text{VIX}_t}{K}\right)\right)^2 + 4\left(\ln\left(\frac{\text{VIX}_t}{K}\right)\right)^3}{K^2}\widehat{P}(t,T,K)dK, \end{aligned} \tag{D.3}$$

where $\widehat{C}(t,T,K)$ and $\widehat{P}(t,T,K)$ are the market prices of OTM VIX call and put options with a maturity of T and a strike price of K. The implied volatility and skewness are then computed by

$$\begin{aligned} \text{IV}(t,T) &= \sqrt{\frac{\exp(r_t(T-t))H_1(t,T) - \mu(t,T)^2}{T-t}} \\ \text{SKEW}(t,T) &= \frac{\exp(r_t(T-t))H_2(t,T) - 3\mu(t,T)\exp(r_t(T-t))H_1(t,T) + 2\mu(t,T)^3}{(T-t)^{3/2}\text{IV}(t,T)^3}, \end{aligned} \tag{D.4}$$

where $\text{IV}(t,T)$ is annualized and

$$\mu(t,T) = \frac{F(t,T)}{\text{VIX}_t} - 1 - \frac{\exp(r_t(T-t))}{2}H_1(t,T) - \frac{\exp(r_t(T-t))}{6}H_2(t,T) - \frac{\exp(r_t(T-t))}{24}H_3(t,T). \tag{D.5}$$

In practice, the computation of the implied volatility and skewness is prone to

truncation and discretization errors, as reported by Jiang and Tian (2007). To reduce such errors, we interpolate or extrapolate option prices on a fine grid of varying strike prices following the approach of Carr and Wu (2009). First, for a given date and a given maturity, we require at least two observations for both puts and calls. We then generate a grid of strike prices with one-point increments, interpolate Black-Scholes implied volatility at each point, and translate the interpolated implied volatility into the option price. In addition, an extrapolation is used for grid points beyond the last observed strike price.

E AVV model with double exponential jumps

Kou (2002) and Kou and Wang (2004) develop a double exponential jump diffusion model in which both upward and downward jumps are driven by a single compound Poisson process. To be specific, we introduce the dynamics for v_t as follows:

$$\begin{aligned} dv_t &= \kappa_v(u_t - v_t)dt + \sqrt{w_t}dB_{1t}^{\mathbb{Q}} + J^{\mathbb{Q}}dN_t^{\mathbb{Q}} - \lambda[p\delta_+ + (1-p)\delta_-]dt \\ du_t &= \kappa_u(\overline{u} - u_t)dt + \sigma_u dB_{2t}^{\mathbb{Q}} \\ dw_t &= \kappa_w(\overline{w} - w_t)dt + \sigma_w\sqrt{w_t}dB_{3t}^{\mathbb{Q}}, \end{aligned} \tag{E.1}$$

where $N_t^{\mathbb{Q}}$ denotes a risk-neutral Poisson process driving jumps with the intensity λ and $J^{\mathbb{Q}}$ has a probability density function that takes $\frac{p}{\delta_+}\exp(-x/\delta_+)$ if $x > 0$ and $\frac{1-p}{|\delta_-|}\exp(-x/\delta_-)$ if $x < 0$. The coefficients of the cumulant generating function of this model satisfy the following system of ODEs:

$$\begin{aligned} \dot{\alpha}(s) &= \kappa_u\overline{u}\beta_u(s) + \kappa_w\overline{w}\beta_w(s) + \frac{1}{2}\sigma_u^2\beta_u(s)^2 \\ &\quad - \lambda\left[p\delta_+ + (1-p)\delta_-\right]\beta_v(s) + \lambda\left[\frac{p}{1-\delta_+\beta_v(s)} + \frac{1-p}{1-\delta_-\beta_v(s)} - 1\right] \\ \dot{\beta}_v(s) &= -\kappa_v\beta_v(s) \\ \dot{\beta}_u(s) &= \kappa_v\beta_v(s) - \kappa_u\beta_u(s) \\ \dot{\beta_w}(s) &= -\kappa_w\beta_w(s) + \frac{1}{2}\beta_v(s)^2 + \frac{1}{2}\sigma_w^2\beta_w(s)^2 + \rho\sigma_w\beta_v(s)\beta_w(s), \end{aligned} \tag{E.2}$$

where the boundary conditions are given as $\alpha(0) = 0$, $\beta_v(0) = \mathrm{i}\phi$, $\beta_u(0) = 0$, and $\beta_w(0) = 0$.

Note that the AVV model with double exponential jumps is different from the

AVV-AJ model where upward and downward jumps are driven by independent compound Poisson processes with each having its own jump intensity and jump-size distribution. In the AVV-AJ model, upward and downward jumps have distinct jump-size distributions so a combination of upward and downward jumps cannot collapse to a single compound Poisson process as in the double exponential jumps.

Table 1: **Summary of model specifications**

Model	Description	Constraints
SVV	Symmetric volatility, no jumps	$\rho = 0$, $\lambda_+ = 0$, and $\lambda_- = 0$
AVV	Asymmetric volatility, no jumps	$\lambda_+ = 0$ and $\lambda_- = 0$
AVV-UJ	Asymmetric volatility, upward jumps only	$\lambda_- = 0$
AVV-AJ	Asymmetric volatility, asymmetric jumps	Not applicable

Table 2: **Summary statistics for VIX futures and options**

The moneyness is defined as $d = K/F(t,T)$, where K is the strike price and $F(t,T)$ is the futures price with a strike price of K and a maturity of T. Short-term contracts are those with no more than three months to maturity, and long-term contracts are those with more than three months to maturity. BSIV stands for Black-Scholes implied volatility. The futures data come from Thomson Reuters, Datastream, while the options data are from OptionMetrics, Ivy DB.

		Months to maturity		
		Short	Long	Total
All futures	No. of futures	4,578	4,752	9,330
	Average price	24.55	25.21	24.89
All OTM options	No. of options	49,992	52,533	102,525
	Average price	1.25	1.86	1.56
	Average BSIV	1.05	0.84	0.94
	Average vega	2.58	4.60	3.61
OTM puts	No. of options	13,758	13,949	27,707
	Average price	1.33	1.90	1.62
	Average BSIV	0.68	0.52	0.60
	Average vega	2.86	5.01	3.94
OTM calls	No. of options	36,403	38,696	75,099
	Average price	1.23	1.85	1.55
	Average BSIV	1.20	0.95	1.07
	Average vega	2.47	4.45	3.49
Deep OTM puts	No. of options	6,382	7,639	14,021
($d < 0.85$)	Average price	0.75	1.14	0.96
	Average BSIV	0.65	0.48	0.56
	Average vega	2.61	4.38	3.58
Moderate OTM puts	No. of options	7,384	6,312	13,696
($0.85 < d < 1.0$)	Average price	1.83	2.82	2.29
	Average BSIV	0.70	0.57	0.64
	Average vega	3.08	5.76	4.31
Moderate OTM calls	No. of options	17,424	17,433	34,857
($1.0 < d < 1.3$)	Average price	1.77	2.68	2.23
	Average BSIV	1.16	0.94	1.05
	Average vega	2.85	5.14	4.00
Deep OTM calls	No. of options	18,987	21,275	40,262
($d > 1.3$)	Average price	0.73	1.16	0.96
	Average BSIV	1.23	0.96	1.09
	Average vega	2.12	3.89	3.06

Table 3: **Parameter estimates: In-sample**

This table shows the parameter estimates that are obtained using the full sample period from July 2006 through January 2013. AIC and SIC stand for Akaike and Schwarz information criteria, respectively. Standard errors are in parentheses.

	MS	SVV	AVV	AVV-UJ	AVV-AJ
κ_v	8.956	6.384	7.393	6.488	6.576
	(0.008)	(0.006)	(0.006)	(0.005)	(0.006)
ρ			0.870	0.422	0.794
			(0.004)	(0.005)	(0.009)
κ_u	0.386	0.324	0.319	0.256	0.258
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
$\overline{u}$	2.979	2.877	2.997	3.073	3.106
	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)
σ_u	0.088	0.442	0.419	0.282	0.293
	(0.004)	(0.001)	(0.002)	(0.003)	(0.003)
η_u	-0.039	-0.003	-0.022	-0.019	-0.024
	(0.013)	(0.013)	(0.027)	(0.016)	(0.017)
κ_w		5.506	1.553	1.050	0.847
		(0.015)	(0.008)	(0.008)	(0.008)
$\overline{w}$		1.613	1.629	2.137	1.956
		(0.003)	(0.012)	(0.019)	(0.027)
σ_w		7.271	2.610	2.379	1.976
		(0.011)	(0.009)	(0.013)	(0.012)
η_w		-28.429	-2.184	-2.926	-2.150
		(1.767)	(0.642)	(1.036)	(0.869)
λ_z	3.358				
	(0.013)				
δ_z	3.391				
	(0.007)				
λ_+				2.664	2.682
				(0.039)	(0.036)
δ_+				0.267	0.266
				(0.001)	(0.001)
λ_-					2.042
					(0.152)
δ_-					-0.217
					(0.009)
σ_e	0.080	0.082	0.078	0.072	0.071
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
σ_f	0.031	0.031	0.030	0.030	0.029
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log(L)	8,071	8,600	15,803	24,642	24,826
AIC	-16,124	-17,177	-31,582	-49,255	-49,620
SIC	-16,075	-17,118	-31,517	-49,179	-49,534

Table 4: **Parameter estimates: Out-of-sample**

This table shows the parameter estimates that are obtained using a subsample of four years from July 2006 through June 2010. This parameter set is used for an out-of-sample test spanning July 2010 through January 2013. AIC and SIC stand for Akaike and Schwarz information criteria, respectively. Standard errors are in parentheses.

	MS	SVV	AVV	AVV-UJ	AVV-AJ
κ_v	11.512	9.602	10.074	9.458	9.652
	(0.014)	(0.013)	(0.011)	(0.010)	(0.011)
ρ			0.606	0.381	0.683
			(0.006)	(0.007)	(0.012)
κ_u	0.495	0.477	0.427	0.399	0.403
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
$\overline{u}$	2.751	2.714	2.782	2.785	2.796
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
σ_u	0.236	0.414	0.402	0.303	0.309
	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)
η_u	0.025	0.019	0.017	0.016	0.014
	(0.021)	(0.025)	(0.050)	(0.014)	(0.017)
κ_w		2.142	1.440	1.087	0.991
		(0.014)	(0.011)	(0.011)	(0.011)
$\overline{w}$		2.709	2.565	3.464	3.146
		(0.021)	(0.027)	(0.042)	(0.047)
σ_w		6.317	3.574	3.067	2.861
		(0.016)	(0.022)	(0.032)	(0.035)
η_w		-10.625	-2.589	-3.395	-3.023
		(2.050)	(1.485)	(1.734)	(1.721)
λ_z	2.022				
	(0.015)				
δ_z	4.482				
	(0.019)				
λ_+				1.759	2.216
				(0.076)	(0.082)
δ_+				0.283	0.265
				(0.005)	(0.004)
λ_-					9.080
					(0.879)
δ_-					-0.120
					(0.006)
σ_e	0.070	0.073	0.069	0.067	0.067
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
σ_f	0.026	0.025	0.024	0.024	0.024
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log(L)	7,086	7,207	10,430	12,129	12,323
AIC	-14,155	-14,392	-20,836	-24,231	-24,613
SIC	-14,110	-14,338	-20,777	-24,162	-24,535

Table 5: **Pricing performance across different models: In-sample**

This table shows the in-sample performance metrics across the different models. The metric RMSE_i stands for the root mean squared error of a model i as given by Equation (16). The metric $\Delta\text{RMSE}_{i|j}$ measures the pricing performance difference of a model i relative to a model j, which is defined as $\Delta\text{RMSE}_{i|j} = 100 \times \log(\text{RMSE}_i/\text{RMSE}_j)$. A negative (positive) value of $\Delta\text{RMSE}_{i|j}$ means that the pricing performance of the model i is better (worse) than that of the model j by a percentage of that value. The moneyness is defined as $d = K/F(t,T)$, where K is the strike price and $F(t,T)$ is the futures price with a strike price of K and a maturity of T. Short-term contracts are those with no more than three months to maturity, and long-term contracts are those with more than three months to maturity.

	RMSE_i			$\Delta\text{RMSE}_{i\|\text{SVV}}$			$\Delta\text{RMSE}_{i\|\text{AVV}}$			$\Delta\text{RMSE}_{i\|\text{AVV-UJ}}$		
	Short	long	Total	Short	Long	Total	Short	Long	Total	Short	Long	Total
Panel A: All futures												
MS	0.1867	0.1839	0.1853	-5.22	8.71	1.18	6.02	4.76	5.39	1.49	10.91	5.93
SVV	0.1916	0.1761	0.1842	0.00	0.00	0.00	11.24	-3.95	4.22	6.71	2.21	4.75
AVV	0.1812	0.1796	0.1804	-11.24	3.95	-4.22	0.00	0.00	0.00	-4.53	6.15	0.53
AVV-UJ	0.1853	0.1741	0.1799	-6.71	-2.21	-4.75	4.53	-6.15	-0.53	0.00	0.00	0.00
AVV-AJ	0.1836	0.1744	0.1791	-8.55	-1.88	-5.60	2.70	-5.83	-1.39	-1.83	0.32	-0.86
Panel B: All options												
MS	0.1218	0.0829	0.1037	8.25	13.81	10.00	13.43	21.30	15.87	20.83	30.61	23.83
SVV	0.1122	0.0722	0.0939	0.00	0.00	0.00	5.18	7.48	5.87	12.59	16.80	13.83
AVV	0.1065	0.0670	0.0885	-5.18	-7.48	-5.87	0.00	0.00	0.00	7.41	9.32	7.96
AVV-UJ	0.0989	0.0611	0.0817	-12.59	-16.80	-13.83	-7.41	-9.32	-7.96	0.00	0.00	0.00
AVV-AJ	0.0985	0.0609	0.0814	-13.01	-17.10	-14.21	-7.82	-9.62	-8.35	-0.42	-0.30	-0.38
Panel C: OTM puts												
MS	0.1331	0.0785	0.1091	-7.91	0.35	-5.88	7.87	11.20	8.72	5.21	12.98	7.12
SVV	0.1440	0.0782	0.1157	0.00	0.00	0.00	15.78	10.85	14.60	13.12	12.63	13.01
AVV	0.1230	0.0702	0.1000	-15.78	-10.85	-14.60	0.00	0.00	0.00	-2.66	1.78	-1.60
AVV-UJ	0.1263	0.0689	0.1016	-13.12	-12.63	-13.01	2.66	-1.78	1.60	0.00	0.00	0.00
AVV-AJ	0.1245	0.0685	0.1003	-14.61	-13.23	-14.29	1.17	-2.37	0.32	-1.49	-0.60	-1.28
Panel D: OTM calls												
MS	0.1183	0.0846	0.1023	18.00	18.93	18.33	16.56	24.92	19.35	30.59	37.72	32.98
SVV	0.0988	0.0700	0.0852	0.00	0.00	0.00	-1.44	5.99	1.02	12.59	18.79	14.66
AVV	0.1002	0.0659	0.0843	1.44	-5.99	-1.02	0.00	0.00	0.00	14.03	12.80	13.64
AVV-UJ	0.0871	0.0580	0.0736	-12.59	-18.79	-14.66	-14.03	-12.80	-13.64	0.00	0.00	0.00
AVV-AJ	0.0874	0.0579	0.0737	-12.22	-18.94	-14.45	-13.66	-12.94	-13.43	0.37	-0.15	0.21
Panel E: Deep OTM puts ($d < 0.85$)												
MS	0.1152	0.0695	0.0932	-7.38	-4.78	-6.61	19.07	9.68	16.03	18.07	10.69	15.72
SVV	0.1241	0.0729	0.0995	0.00	0.00	0.00	26.45	14.46	22.64	25.46	15.48	22.33
AVV	0.0952	0.0631	0.0794	-26.45	-14.46	-22.64	0.00	0.00	0.00	-0.99	1.01	-0.31
AVV-UJ	0.0962	0.0625	0.0796	-25.46	-15.48	-22.33	0.99	-1.01	0.31	0.00	0.00	0.00
AVV-AJ	0.0923	0.0617	0.0772	-29.54	-16.69	-25.43	-3.09	-2.23	-2.79	-4.09	-1.21	-3.10
Panel F: Moderate OTM puts ($0.85 < d < 1.0$)												
MS	0.1467	0.0881	0.1232	-8.19	4.59	-5.46	2.79	12.38	4.89	-0.50	14.77	2.70
SVV	0.1592	0.0842	0.1301	0.00	0.00	0.00	10.98	7.78	10.35	7.69	10.18	8.16
AVV	0.1427	0.0779	0.1173	-10.98	-7.78	-10.35	0.00	0.00	0.00	-3.29	2.39	-2.19
AVV-UJ	0.1475	0.0760	0.1199	-7.69	-10.18	-8.16	3.29	-2.39	2.19	0.00	0.00	0.00
AVV-AJ	0.1466	0.0760	0.1194	-8.25	-10.28	-8.64	2.72	-2.49	1.71	-0.57	-0.10	-0.48
Panel G: Moderate OTM calls ($1.0 < d < 1.3$)												
MS	0.1260	0.1017	0.1145	10.46	24.75	15.62	8.73	30.78	16.32	17.27	45.14	26.50
SVV	0.1135	0.0794	0.0979	0.00	0.00	0.00	-1.73	6.03	0.69	6.81	20.39	10.88
AVV	0.1154	0.0748	0.0972	1.73	-6.03	-0.69	0.00	0.00	0.00	8.54	14.36	10.19
AVV-UJ	0.1060	0.0648	0.0878	-6.81	-20.39	-10.88	-8.54	-14.36	-10.19	0.00	0.00	0.00
AVV-AJ	0.1066	0.0647	0.0882	-6.27	-20.42	-10.50	-8.00	-14.39	-9.81	0.53	-0.03	0.38
Panel H: Deep OTM calls ($d > 1.3$)												
MS	0.1107	0.0674	0.0904	28.76	9.60	22.34	27.82	15.54	23.90	53.14	26.26	43.67
SVV	0.0831	0.0612	0.0723	0.00	0.00	0.00	-0.93	5.94	1.56	24.39	16.66	21.32
AVV	0.0838	0.0577	0.0712	0.93	-5.94	-1.56	0.00	0.00	0.00	25.32	10.72	19.77
AVV-UJ	0.0651	0.0518	0.0584	-24.39	-16.66	-21.32	-25.32	-10.72	-19.77	0.00	0.00	0.00
AVV-AJ	0.0651	0.0516	0.0584	-24.42	-16.96	-21.47	-25.36	-11.02	-19.91	-0.03	-0.30	-0.14

Table 6: **Pricing performance across different models: Out-of-sample**

This table shows the out-of-sample performance metrics across the different models. The metric RMSE_i stands for the root mean squared error of a model i as given by Equation (16). The metric $\Delta\text{RMSE}_{i|j}$ measures the pricing performance difference of a model i relative to a model j, which is defined as $\Delta\text{RMSE}_{i|j} = 100 \times \log(\text{RMSE}_i/\text{RMSE}_j)$. A negative (positive) value of $\Delta\text{RMSE}_{i|j}$ means that the pricing performance of the model i is better (worse) than that of the model j by a percentage of that value. The moneyness is defined as $d = K/F(t,T)$, where K is the strike price and $F(t,T)$ is the futures price with a strike price of K and a maturity of T. Short-term contracts are those with no more than three months to maturity, and long-term contracts are those with more than three months to maturity.

	RMSE_i			$\Delta\text{RMSE}_{i\|\text{SVV}}$			$\Delta\text{RMSE}_{i\|\text{AVV}}$			$\Delta\text{RMSE}_{i\|\text{AVV-UJ}}$		
	Short	long	Total	Short	Long	Total	Short	Long	Total	Short	Long	Total
Panel A: All futures												
MS	0.2415	0.2242	0.2332	1.12	16.03	7.02	14.62	7.88	11.60	13.26	17.64	15.10
SVV	0.2402	0.2069	0.2252	0.00	0.00	0.00	13.50	-8.15	4.58	12.14	1.61	8.08
AVV	0.2245	0.2155	0.2201	-13.50	8.15	-4.58	0.00	0.00	0.00	-1.36	9.76	3.51
AVV-UJ	0.2261	0.2052	0.2163	-12.14	-1.61	-8.08	1.36	-9.76	-3.51	0.00	0.00	0.00
AVV-AJ	0.2251	0.2073	0.2166	-13.02	0.34	-7.78	0.48	-7.81	-3.21	-0.89	1.95	0.30
Panel B: All options												
MS	0.1625	0.1320	0.1475	15.87	28.46	20.71	19.95	31.97	24.58	24.71	38.34	29.91
SVV	0.1386	0.0993	0.1199	0.00	0.00	0.00	4.07	3.51	3.87	8.84	9.88	9.20
AVV	0.1331	0.0959	0.1154	-4.07	-3.51	-3.87	0.00	0.00	0.00	4.76	6.37	5.33
AVV-UJ	0.1269	0.0900	0.1094	-8.84	-9.88	-9.20	-4.76	-6.37	-5.33	0.00	0.00	0.00
AVV-AJ	0.1263	0.0911	0.1095	-9.30	-8.56	-9.04	-5.23	-5.05	-5.16	-0.47	1.32	0.16
Panel C: OTM puts												
MS	0.1728	0.0794	0.1342	8.05	-1.02	6.33	20.56	6.49	17.78	21.87	10.24	19.62
SVV	0.1594	0.0802	0.1260	0.00	0.00	0.00	12.51	7.51	11.45	13.82	11.25	13.29
AVV	0.1407	0.0744	0.1123	-12.51	-7.51	-11.45	0.00	0.00	0.00	1.31	3.74	1.84
AVV-UJ	0.1388	0.0716	0.1103	-13.82	-11.25	-13.29	-1.31	-3.74	-1.84	0.00	0.00	0.00
AVV-AJ	0.1361	0.0719	0.1087	-15.78	-10.93	-14.76	-3.27	-3.42	-3.30	-1.96	0.32	-1.47
Panel D: OTM calls												
MS	0.1612	0.1440	0.1525	18.53	32.18	24.41	19.77	35.01	26.27	25.69	41.84	32.55
SVV	0.1339	0.1044	0.1194	0.00	0.00	0.00	1.24	2.83	1.86	7.16	9.66	8.14
AVV	0.1323	0.1015	0.1172	-1.24	-2.83	-1.86	0.00	0.00	0.00	5.93	6.83	6.28
AVV-UJ	0.1246	0.0948	0.1101	-7.16	-9.66	-8.14	-5.93	-6.83	-6.28	0.00	0.00	0.00
AVV-AJ	0.1247	0.0962	0.1108	-7.10	-8.17	-7.52	-5.86	-5.35	-5.66	0.06	1.49	0.62
Panel E: Deep OTM puts ($d < 0.85$)												
MS	0.0896	0.0628	0.0754	-13.83	-7.55	-11.41	3.23	0.46	2.11	7.15	2.71	5.33
SVV	0.1029	0.0678	0.0845	0.00	0.00	0.00	17.05	8.01	13.52	20.98	10.25	16.74
AVV	0.0867	0.0626	0.0738	-17.05	-8.01	-13.52	0.00	0.00	0.00	3.92	2.24	3.22
AVV-UJ	0.0834	0.0612	0.0715	-20.98	-10.25	-16.74	-3.92	-2.24	-3.22	0.00	0.00	0.00
AVV-AJ	0.0803	0.0613	0.0700	-24.78	-10.12	-18.84	-7.72	-2.10	-5.33	-3.80	0.14	-2.10
Panel F: Moderate OTM puts ($0.85 < d < 1.0$)												
MS	0.2309	0.1079	0.1924	12.72	4.69	11.66	24.05	11.70	22.36	24.74	16.94	23.71
SVV	0.2033	0.1030	0.1712	0.00	0.00	0.00	11.33	7.01	10.69	12.02	12.25	12.05
AVV	0.1815	0.0960	0.1539	-11.33	-7.01	-10.69	0.00	0.00	0.00	0.69	5.24	1.36
AVV-UJ	0.1803	0.0911	0.1518	-12.02	-12.25	-12.05	-0.69	-5.24	-1.36	0.00	0.00	0.00
AVV-AJ	0.1775	0.0916	0.1499	-13.56	-11.74	-13.30	-2.24	-4.73	-2.61	-1.55	0.51	-1.25
Panel G: Moderate OTM calls ($1.0 < d < 1.3$)												
MS	0.1817	0.1976	0.1897	-0.07	32.47	14.69	0.10	35.70	15.98	4.78	46.37	22.74
SVV	0.1819	0.1428	0.1638	0.00	0.00	0.00	0.16	3.23	1.29	4.85	13.90	8.05
AVV	0.1816	0.1383	0.1617	-0.16	-3.23	-1.29	0.00	0.00	0.00	4.69	10.67	6.76
AVV-UJ	0.1732	0.1243	0.1511	-4.85	-13.90	-8.05	-4.69	-10.67	-6.76	0.00	0.00	0.00
AVV-AJ	0.1733	0.1266	0.1521	-4.85	-12.07	-7.43	-4.68	-8.83	-6.14	0.01	1.84	0.62
Panel H: Deep OTM calls ($d > 1.3$)												
MS	0.1448	0.0996	0.1225	53.73	31.53	44.65	58.67	33.44	48.17	69.36	32.68	53.06
SVV	0.0846	0.0727	0.0784	0.00	0.00	0.00	4.94	1.92	3.52	15.63	1.15	8.40
AVV	0.0805	0.0713	0.0757	-4.94	-1.92	-3.52	0.00	0.00	0.00	10.69	-0.77	4.88
AVV-UJ	0.0724	0.0718	0.0721	-15.63	-1.15	-8.40	-10.69	0.77	-4.88	0.00	0.00	0.00
AVV-AJ	0.0726	0.0724	0.0725	-15.33	-0.28	-7.80	-10.39	1.63	-4.28	0.29	0.86	0.60

Table 7: **Pairwise model comparisons: In-sample**

This table shows the Diebold and Mariano (2002) test statistics of the in-sample performance differences. A negative (positive) statistic in a cell (i, j) indicates that the model i outperforms (underperforms) the model j. The symbols *, **, and *** represent the statistical significance of the performance difference at the 10, 5, and 1 percent levels, respectively.

	MS	SVV	AVV	AVV-UJ	AVV-AJ
			Panel A: Futures		
MS	0.00	0.35	3.20***	1.96*	2.56**
SVV	-0.35	0.00	2.52**	2.18**	2.79***
AVV	-3.20***	-2.52**	0.00	-0.50	0.60
AVV-UJ	-1.96*	-2.18**	0.50	0.00	4.93***
AVV-AJ	-2.56**	-2.79***	-0.60	-4.93***	0.00
			Panel B: Options		
MS	0.00	6.44***	11.50***	11.82***	11.76***
SVV	-6.44***	0.00	7.08***	11.22***	10.72***
AVV	-11.50***	-7.08***	0.00	8.16***	8.44***
AVV-UJ	-11.82***	-11.22***	-8.16***	0.00	1.45
AVV-AJ	-11.76***	-10.72***	-8.44***	-1.45	0.00

Table 8: **Pairwise model comparisons: Out-of-sample**

This table shows the Diebold and Mariano (2002) test statistics of the out-of-sample performance differences. A negative (positive) statistic in a cell (i, j) indicates that the model i outperforms (underperforms) the model j. The symbols *, **, and *** represent the statistical significance of the performance difference at the 10, 5, and 1 percent levels, respectively.

	MS	SVV	AVV	AVV-UJ	AVV-AJ
			Panel A: Futures		
MS	0.00	7.58***	14.25***	17.46***	17.15***
SVV	-7.58***	0.00	8.17***	15.77***	15.51***
AVV	-14.25***	-8.17***	0.00	6.50***	8.27***
AVV-UJ	-17.46***	-15.77***	-6.50***	0.00	-1.16
AVV-AJ	-17.15***	-15.51***	-8.27***	1.16	0.00
			Panel B: Options		
MS	0.00	12.26***	14.01***	15.07***	14.96***
SVV	-12.26***	0.00	6.49***	11.87***	10.79***
AVV	-14.01***	-6.49***	0.00	11.71***	12.10***
AVV-UJ	-15.07***	-11.87***	-11.71***	0.00	-1.13
AVV-AJ	-14.96***	-10.79***	-12.10***	1.13	0.00

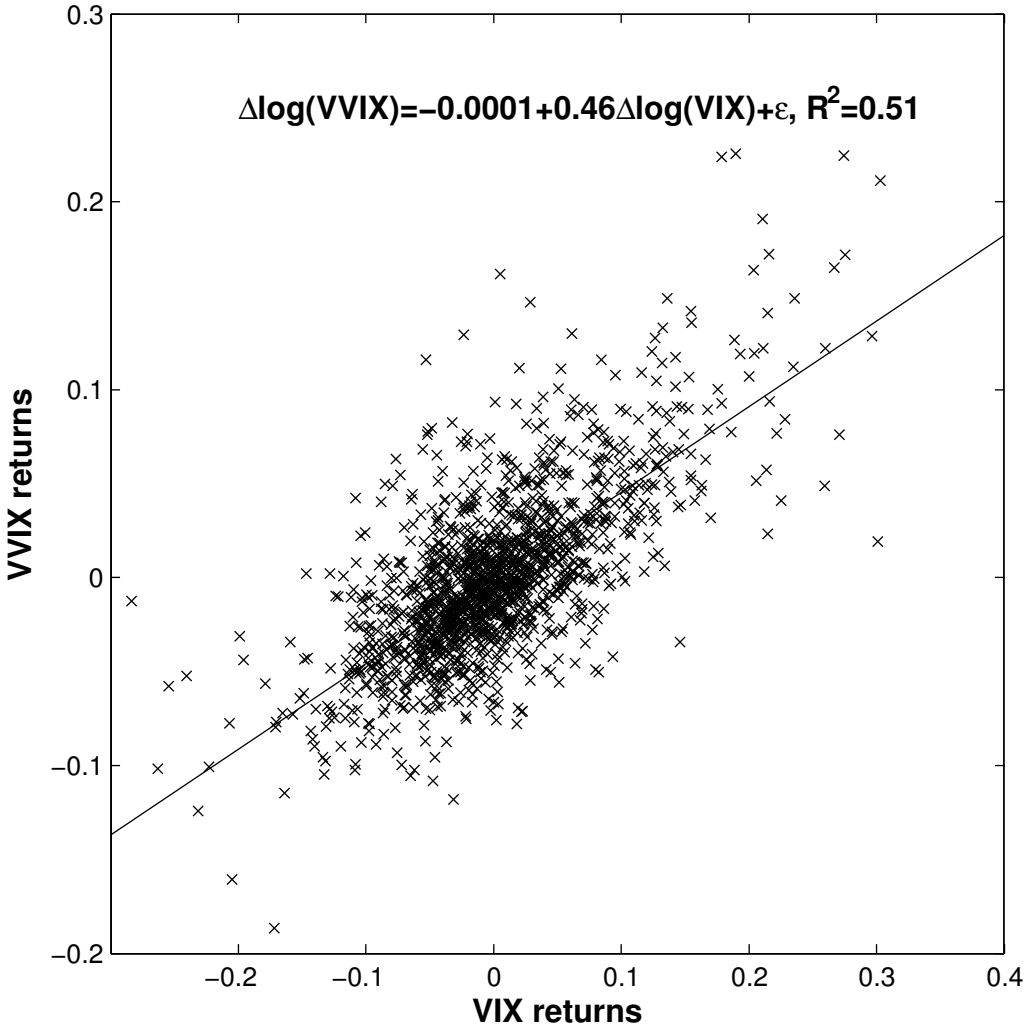

Figure 1: Scatter plot of VVIX log returns against VIX log returns.

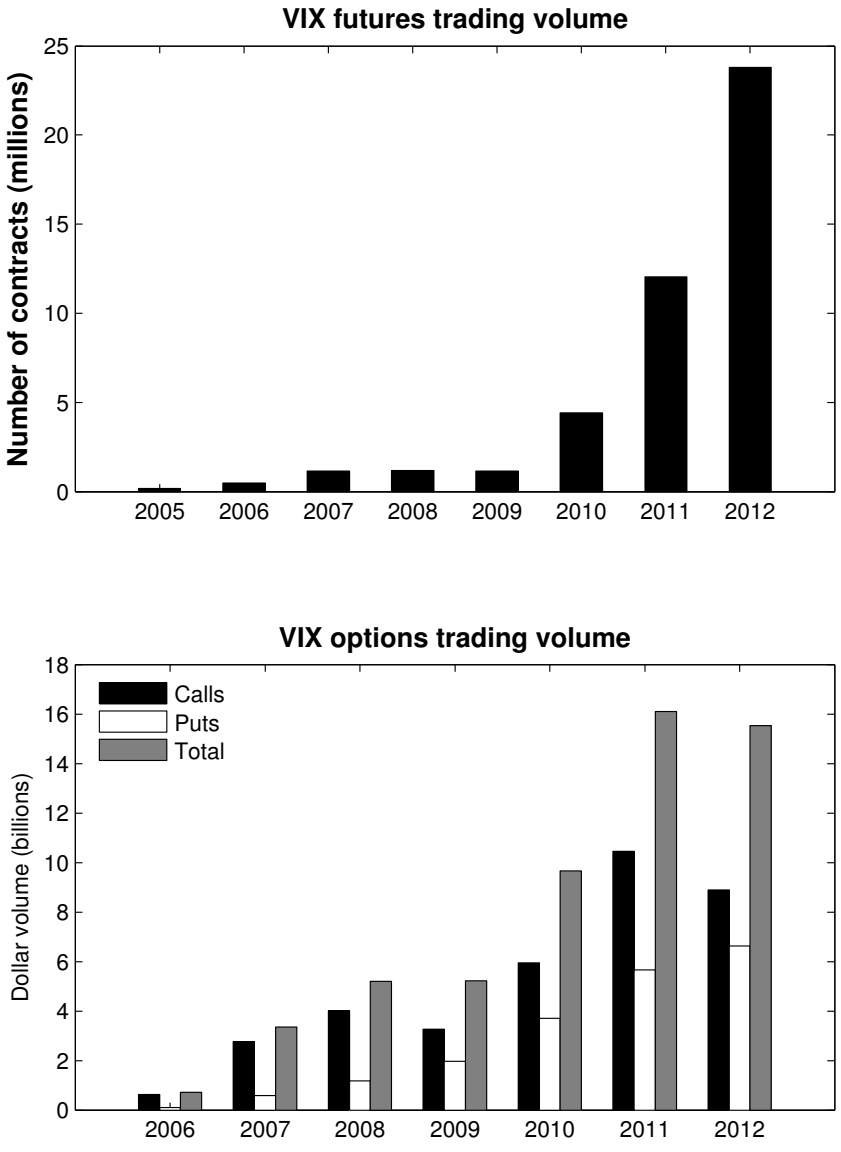

Figure 2: Trading volume of VIX futures and options.
The top panel shows the numbers of futures contracts traded over time, and the bottom panel presents the dollar trading volumes for options contracts over time.

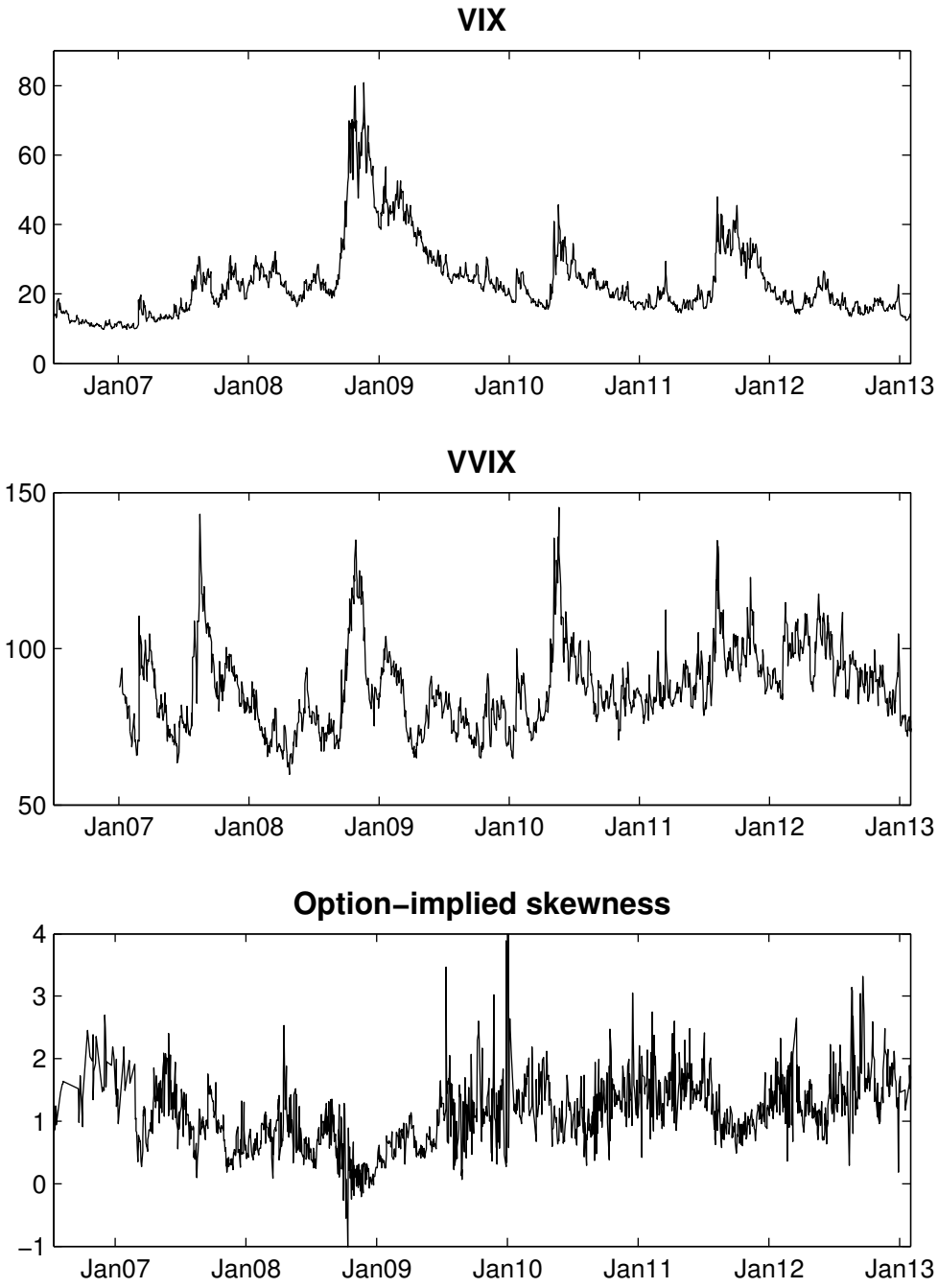

Figure 3: Time series of the VIX, the VVIX, and the option-implied skewness.
The VIX and the VVIX are provided by the CBOE, and the skewness is computed using the model-free approach of Bakshi, Kapadia, and Madan (2003) with a modification. The options data are from OptionMetrics, Ivy DB.

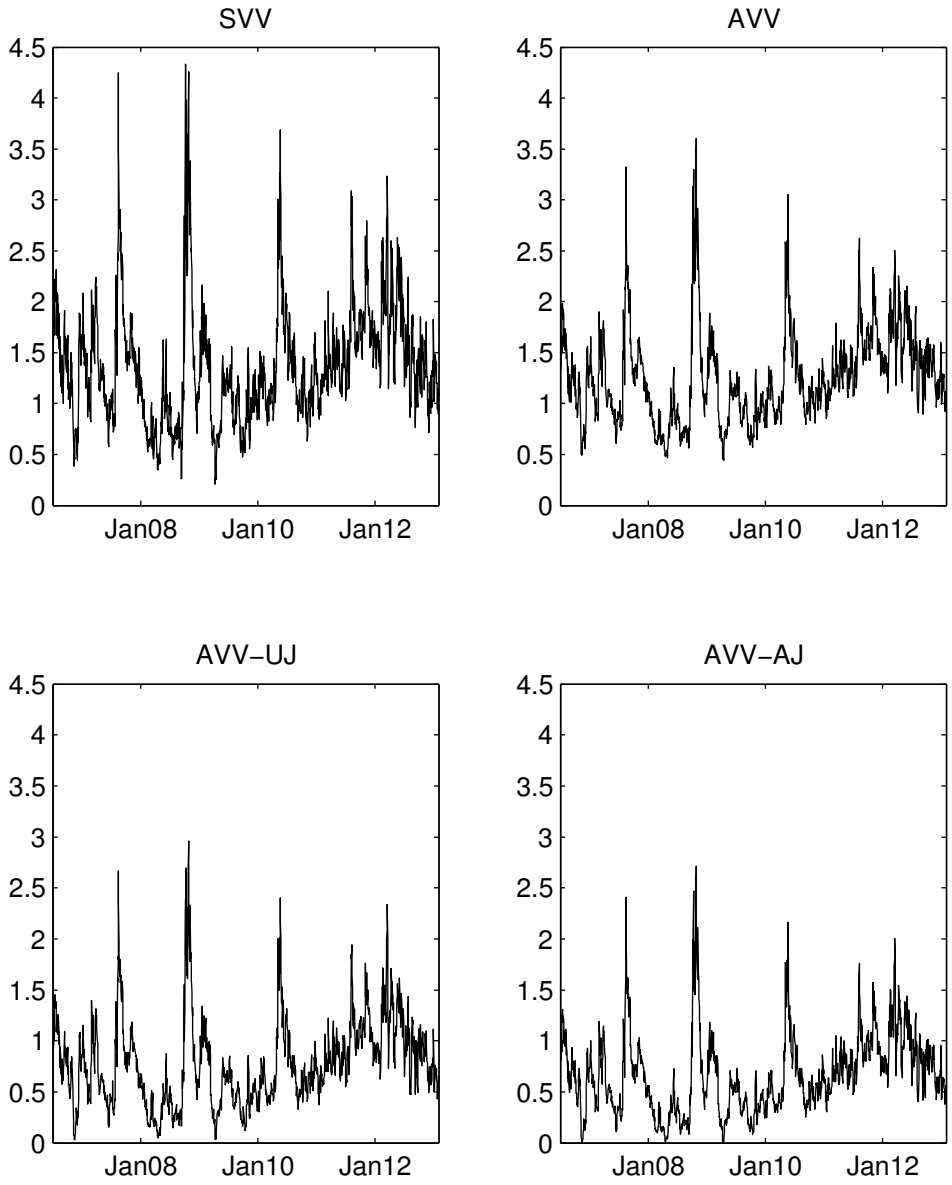

Figure 4: Filtered volatility states across different models.
Each panel displays a time series of the filtered volatility states that are obtained by applying an extended Kalman filter to the sample data from July 2006 through January 2013.

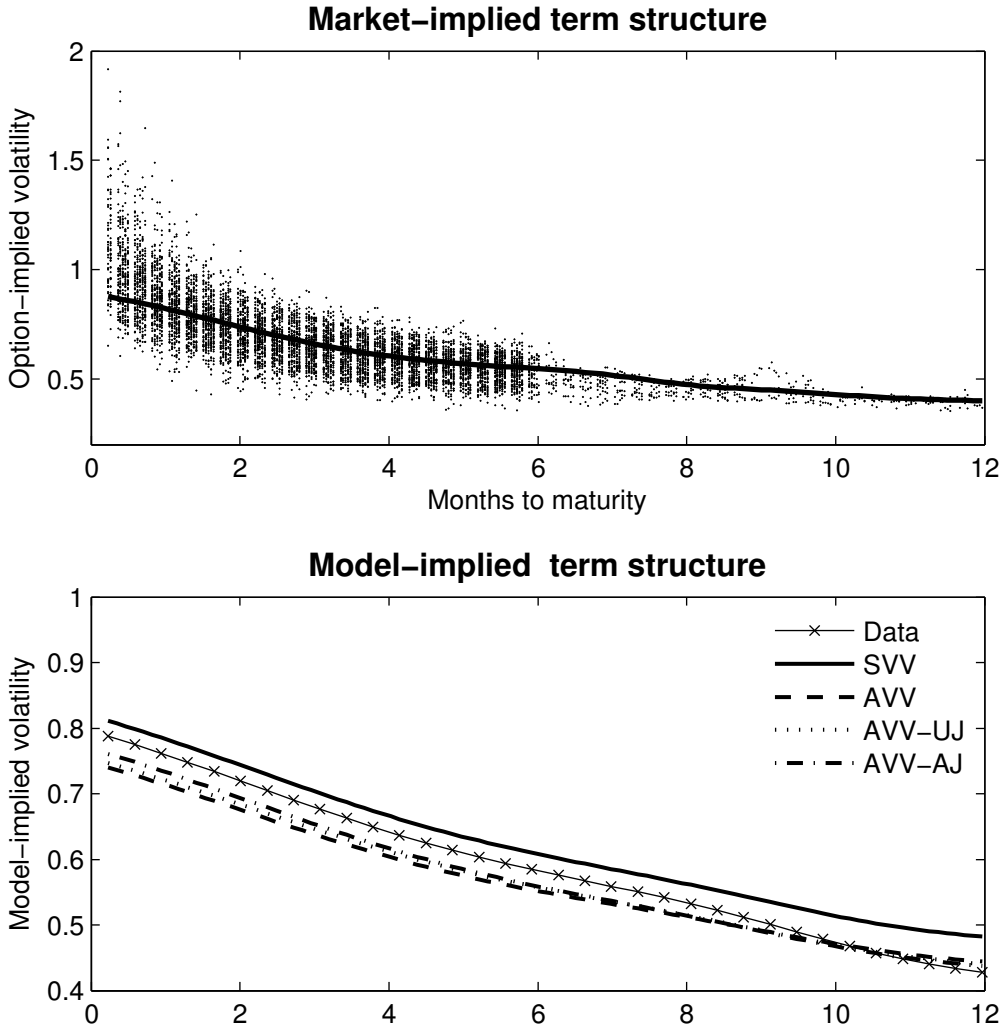

Figure 5: Term structure of volatility.
The top panel shows the scatter plot and the average term structure of the market-implied volatility, and the bottom panel shows the average term structures of the model-implied volatility.

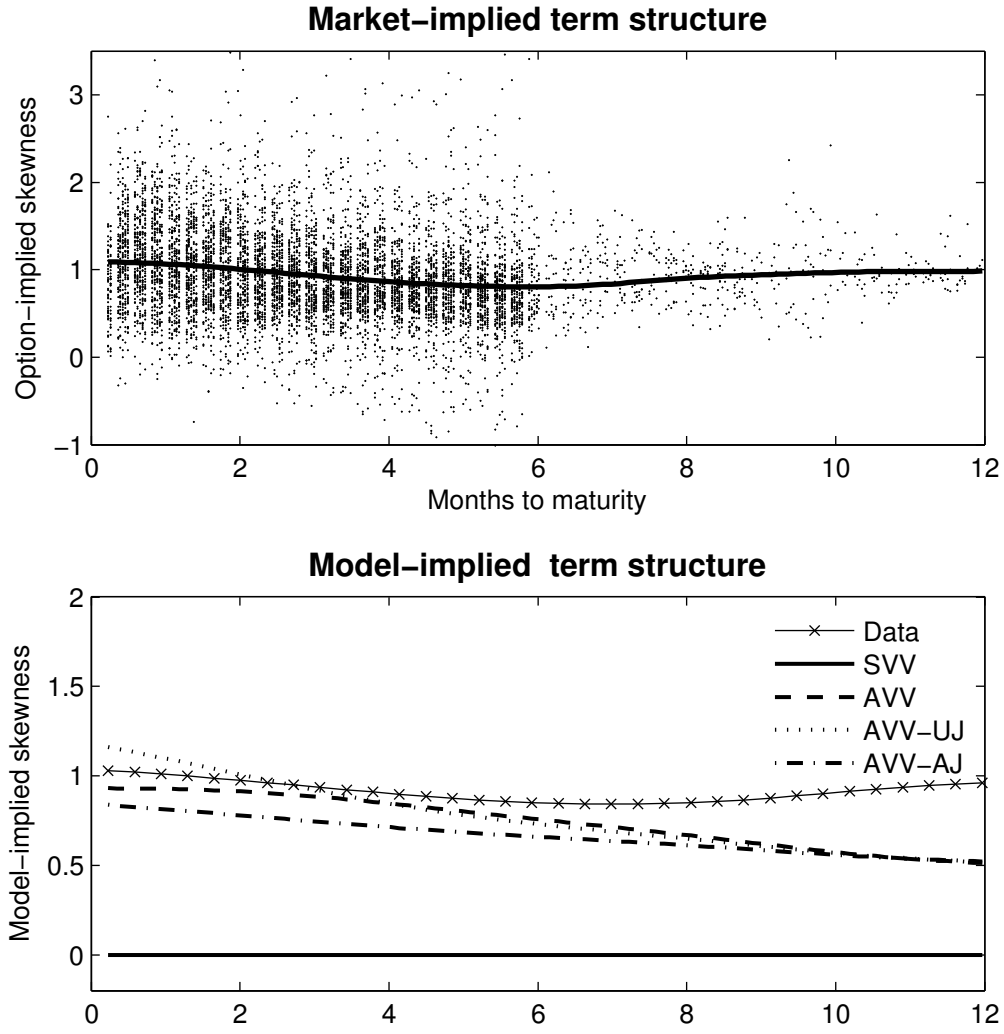

Figure 6: Term structure of skewness.
The top panel shows the scatter plot and the average term structure of the market-implied skewness, and the bottom panel shows the average term structures of the model-implied skewness.