Received: 5 July 2017 | Revised: 22 March 2018 | Accepted: 1 April 2018

DOI: 10.1002/fut.21925

WILEY FUTURES MARKETS

RESEARCH ARTICLE

VIX futures pricing with conditional skewness

Xinglin Yang | **Peng Wang**

Institute of Chinese Financial Studies, Southwestern University of Finance and Economics, Chengdu, China

Correspondence
Xinglin Yang, Institute of Chinese Financial Studies, Southwestern University of Finance and Economics, 555 Liutai Avenue, 611130 Chengdu, Sichuan, China.
Email: xinglinyang@126.com

Funding information
National Natural Science Foundation of China, Grant/Award Number: 71473200

We develop a closed-form VIX futures valuation formula based on the inverse Gaussian GARCH process by Christoffersen et al. that combines conditional skewness, conditional heteroskedasticity, and a leverage effect. The new model outperforms the benchmark in fitting the S&P 500 returns and the VIX futures prices. The fat-tailed innovation underlying the model substantially reduced pricing errors during the 2008 financial crisis. The in- and out-of-sample pricing performance indicates that the new model should be a default modeling choice for pricing the medium- and long-term VIX futures.

KEYWORDS
closed-form formula, conditional skewness, GARCH, pricing kernel, VIX futures

JEL CLASSIFICATION
G13

1 | INTRODUCTION

The VIX, introduced by the Chicago Board Options Exchange (CBOE) in 1993, is an important risk indicator. It is often referred to as the "fear index" by leading financial publications and business news shows. The purpose of the VIX is to measure S&P 500 volatility over the next month as implied by stock index option prices. The VIX has become a particularly popular measure of the U.S. stock market volatility. CBOE introduced exchange-traded VIX futures and options as tradable assets in March 2004 and February 2006, respectively. In the decade since their launch, their combined trading activity has grown markedly, now totaling approximately 800,000 contracts per day.

The fast development of the financial market has increased the importance of speculation and hedging against unexpected changes in volatility. The crucial roles of VIX derivatives in such hedging have been studied widely in the academic literature; see, for example, Grünbichler and Longstaff (1996), Dotsis, Psychoyios, and Skiadopoulos (2007), and Mencía and Sentana (2013). To obtain protection from volatility risk, strategies for managing the risk should be built on reliable VIX futures and option valuation models that describe the empirical properties of asset returns adequately.

Compared with the approaches of options valuation, the literature still offers relatively few pricing formulas of futures valuation. However, several meaningful valuation models for VIX futures have been proposed. Zhang and Zhu (2006) price VIX futures assuming stochastic instantaneous variance in the diffusion process of the S&P 500 returns. Considering the importance of jump diffusion, Lin (2007) constructs a valuation model of VIX futures with simultaneous jumps in the price and variance processes. Zhu and Lian (2012), combining stochastic volatility and random jumps, derive a closed-form and exact formula to evaluate VIX futures. To avoid pricing distortions during the financial crisis, Mencía and Sentana (2013) find that a VIX process combining central tendency and stochastic volatility has a better pricing performance for VIX futures and options.

The GARCH process can simultaneously capture the correlation of volatility with returns and the path dependence in volatility, and it has a more parsimonious description in many situations (Bollerslev, 1986; Heston & Nandi, 2000). Compared with continuous models, GARCH-type models readily compute volatility through historical asset prices and

conveniently implement maximum likelihood estimation (MLE). Therefore, GARCH-type models have been used extensively for options valuation (Christoffersen, Feunou, Jacobs, & Meddahi, 2014; Christoffersen, Heston, & Jacobs, 2006; Dorion, 2016; Heston & Nandi, 2000). To the best of our knowledge, only Wang, Shen, Jiang, and Huang (2017) have proposed a closed-form pricing formula for VIX futures based on the Heston–Nandi (HN) GARCH process.

It is worth emphasizing the importance of capturing nonnormality and asymmetry in economics and finance, which has been widely noted (Patton, 2004, 2006; Patton & Timmermann, 2007). However, under the local risk-neutral valuation relationship (RNVR) framework proposed by Duan (1995), the existing GARCH-type models with Gaussian innovations may not sufficiently allow for all the mass in the tails and the asymmetry in the asset return distribution.

Motivated by the stylized fact of conditional skewness in asset returns, we propose a closed-form formula for pricing VIX futures using the discrete-time IG GARCH model in which the innovation follows an inverse Gaussian distribution with a nonzero third moment. The IG GARCH process reinforces the pricing accuracy for the VIX futures through capturing the conditional skewness of the S&P 500 returns. Unlike the approach that obtains risk-neutral dynamics by directly assuming the form of the state-price density, our framework uses only the no-arbitrage assumption and some technical conditions on the investment strategies, without attempting to characterize the preferences underlying the RNVR or fully describe the economic environment. Following the theoretical framework by Christoffersen, Elkamhi, Feunou, and Jacobs (2010), first we propose a pricing kernel to calculate the equivalent martingale measure (EMM) coefficient. Then, we derive the risk-neutral dynamics of the inverse Gaussian process using an EMM coefficient that fills the gap between the physical and the risk-neutral measures.

We test the performance of the IG GARCH model and that of the most related benchmark, the HN GARCH model, using the daily S&P 500 returns and the VIX futures data. First, we estimate the models by fitting the S&P 500 returns using the maximum likelihood method. The empirical results show that the IG GARCH model outperforms the HN GARCH model in fitting the returns, in forecasting 1-day and 1-week volatility, and in absorbing the heteroskedasticity of returns. Second, we optimize the likelihood function defined on the VIX futures errors and analyze the out-of-sample performance. We find that, overall, the IG GARCH model that incorporates conditional skewness has a better in- and out-of-sample VIX futures valuation performance than the HN GARCH model. When we further dissect the pricing performance of the VIX futures across maturity, bias, and VIX level, the IG GARCH model outperforms the HN GARCH model for small bias, for high VIX level, and particularly for medium- and long-term VIX futures contracts. Using the futures contracts of 2017 as an out-of-sample experiment, the IG GARCH model has futures pricing errors that are 19% below those of the benchmark model for medium- and long-term VIX futures contracts. The empirical findings in the joint subsamples that lie at the intersection of bias, maturity, and VIX levels indicate that the IG GARCH model is significantly superior to the HN GARCH model in all categories. Third, the fat-tailed innovation of the IG GARCH model offers an improvement of 11.78% on average on the extreme days of the 2008 financial crisis. Fourth, we jointly fit the S&P 500 returns and the VIX futures. The joint results once again reconfirm that the IG GARCH model with the nonnormal innovation reduces the pricing errors of the benchmark model with the normal innovation.

The remainder of the paper is organized as follows: Section 2 introduces the inverse Gaussian GARCH process and the nested HN GARCH process; Section 3 provides the estimation results on daily returns; Section 4 develops the theoretical framework of risk neutralization and the closed-form formula for VIX futures valuation; Section 5 analyzes the in- and out-of-sample pricing performance with the volatility index and the VIX futures data; Section 6 estimates the models jointly on returns and futures; and Section 7 concludes the paper.

2 | THE MODEL

2.1 | The return process

The daily returns R_{t+1} and the conditional variance h_{t+1} specified by the dynamic inverse Gaussian GARCH model are:

$$R_{t+1} \equiv \log\left(\frac{S_{t+1}}{S_t}\right) = r + \zeta h_{t+1} + \eta y_{t+1}, \tag{1}$$

$$h_{t+1} = w + bh_t + cy_t + \frac{ah_t^2}{y_t}, \tag{2}$$

where S_t is the closing price of the underlying asset at time t, and r is the risk-free rate. Shocks to returns are y_{t+1}, which is assumed to be an independent distributed inverse Gaussian distribution with degree of freedom parameters $\frac{h_{t+1}}{\eta^2}$ and $y_{t+1} \sim \mathrm{IG}(\frac{h_{t+1}}{\eta^2})$, where IG is the inverse Gaussian distribution, and η is a scale parameter.

As η approaches zero, the inverse Gaussian distribution will converge to normal distribution. Then, the Heston and Nandi (2000) GARCH model can be nested in the IG GARCH model, with the following parameterization:

$$v = \lambda - \frac{1}{\eta},\ w = \omega,\ a = \frac{\alpha}{\eta^4},\ b = \beta + \alpha\gamma^2 - \frac{2\alpha}{\eta^2} + \frac{2\alpha\gamma}{\eta},\ c = \alpha - 2\eta\alpha\gamma. \tag{3}$$

The HN GARCH model is

$$R_{t+1} = r + \lambda h_{t+1} + \sqrt{h_{t+1}} z_{t+1}, \tag{4}$$

$$h_{t+1} = \omega + \beta h_t + \alpha(z_t - \gamma\sqrt{h_t})^2, \tag{5}$$

where z_{t+1} follows a standard normal distribution.

2.2 | The conditional moments of the IG GARCH process

The conditional moments of returns can conveniently be derived using the moments of the inverse Gaussian variable.

$$E_t(R_{t+1}) = r + \left(\zeta + \frac{1}{\eta}\right)h_{t+1}, \tag{6}$$

$$\mathrm{Var}_t(R_{t+1}) = h_{t+1}, \tag{7}$$

$$\mathrm{Skew}_t(R_{t+1}) = \frac{3\eta}{\sqrt{h_{t+1}}}, \tag{8}$$

$$\mathrm{Kurt}_t(R_{t+1}) = \frac{15\eta^2}{h_{t+1}}. \tag{9}$$

The variance persistence can be measured by taking the conditional expectation for Equation (2). For $k \geq 1$

$$E_t[h_{t+k}] = w + \eta^4 a + \left(an^2 + b + \frac{c}{\eta^2}\right)E_t[h_{t+k-1}]. \tag{10}$$

The coefficient $an^2 + b + \frac{c}{\eta^2}$ denotes the variance persistence. The corresponding unconditional expectation of variance is $\frac{w + \eta^4 a}{1 - a\eta^2 - b - \frac{c}{\eta^2}}$.

Meanwhile, the leverage covariance of the IG GARCH model can be computed by

$$\mathrm{Cov}_t\left[\log\frac{S_{t+1}}{S_t}, h_{t+2}\right] = \left(\frac{c}{\eta} - \eta^3 a\right)h_{t+1}. \tag{11}$$

According to Equations (4) and (5), the conditional mean and variance of the HN GARCH process can be derived as

$$E_t(R_{t+1}) = r + \lambda h_{t+1} \quad \text{and} \quad \mathrm{Var}_t(R_{t+1}) = h_{t+1}. \tag{12}$$

The persistence can be measured by the coefficient $\beta + \alpha\gamma^2$, the unconditional expectation of variance is $\frac{\omega + \alpha}{1 - \beta - \alpha\gamma^2}$, and the leverage effect is the conditional covariance between returns and variance: $-2\alpha\gamma h_{t+1}$. These moments will be used to compare models in the empirical study.

TABLE 1 Summary statistics for S&P 500 returns, VIX, and 3-month Treasury bills

	S&P 500 returns	VIX	3-Month Treasury bills
Mean	0.000320	19.152	4.780
Median	0.000428	16.230	4.890
Maximum	0.116	80.860	17.140
Minimum	−0.205	9.890	−0.020
Standard deviation	0.010	9.159	3.201
Skewness	−0.626	2.577	0.616
Kurtosis	20.693	8.796	0.821
Jarque-Bera	245,664.4***	13,914.12***	1,242.929***
Observations	13,719	3,213	13,619

Note. This table shows the summary statistics of S&P 500 returns, VIX, and 3-month Treasury bills. Kurtosis denotes the excess kurtosis.
***Significant at 1% level.

3 | DAILY RETURN EMPIRICAL ANALYSIS

In this section, we present the data and estimate the models described in Section 2. Using returns data, we estimate the parameter values for the IG GARCH model and for the benchmark HN GARCH model.

3.1 | Data

We estimate the models using daily close-to-close S&P 500 returns and the time-series of 3-month Treasury bills as the proxy for the risk-free rate from July 3, 1962 to December 30, 2016, yielding 13,719 and 13,619 observations, respectively.[1] Both are obtained from the Center for Research in Security Prices. The VIX index values are collected from CBOE for the period from March 26, 2004 to December 28, 2016. We use a longer time-series of S&P 500 returns than that of the VIX index and futures prices because it is difficult to estimate return dynamics accurately with a short time-series. Table 1 shows the summary statistics for daily S&P 500 log-returns, the VIX index, and 3-month Treasury bills. Note that the skewness in returns is −0.626, obviously negative.

At the same time, we can easily find the heteroskedasticity and leverage effect in the S&P 500 returns from Graph A of Figure 1. To assess the dynamic properties of the S&P 500 returns, we also plot the autocorrelations of returns and squared returns for lags 1 through 40 in Graphs B and C of Figure 1. The horizontal lines represent a two-standard-error confidence around zero. The Graphs show that returns have little persistence, but squared returns are distinctly persistent.

3.2 | MLE on returns

We estimate the models using the maximum likelihood estimate (MLE). Given the conditional information at time t, the conditional density of the IG GARCH model returns at time $t+1$ is

$$\begin{aligned} f_t(R_{t+1}) &= f_t\left(\frac{R_{t+1}-r-\zeta h_{t+1}}{\eta};\frac{h_{t+1}}{\eta^2}\right)\left|\frac{\partial y_{t+1}}{\partial R_{t+1}}\right| \\ &= \frac{h_{t+1}}{\eta^2\sqrt{2\pi\left(\left(R_{t+1}-r-\zeta h_{t+1}\right)/\eta\right)^3}}\exp\left[-\frac{1}{2}\left(\sqrt{\left(R_{t+1}-r-\zeta h_{t+1}\right)/\eta}-\frac{h_{t+1}}{\eta^2\sqrt{\left(R_{t+1}-r-\zeta h_{t+1}\right)/\eta}}\right)^2\right]\left|\frac{1}{\eta}\right|. \end{aligned} \tag{13}$$

Taking the log and summing the log likelihoods over all the observations yield the sample return likelihood function

$$L_{\text{returns}} = \sum_{t=1}^{T-1}\ln f_t(R_{t+1}) \tag{14}$$

[1]If the 3-month Treasury bill is N.A., we replace it with the previous day's risk-free rate.

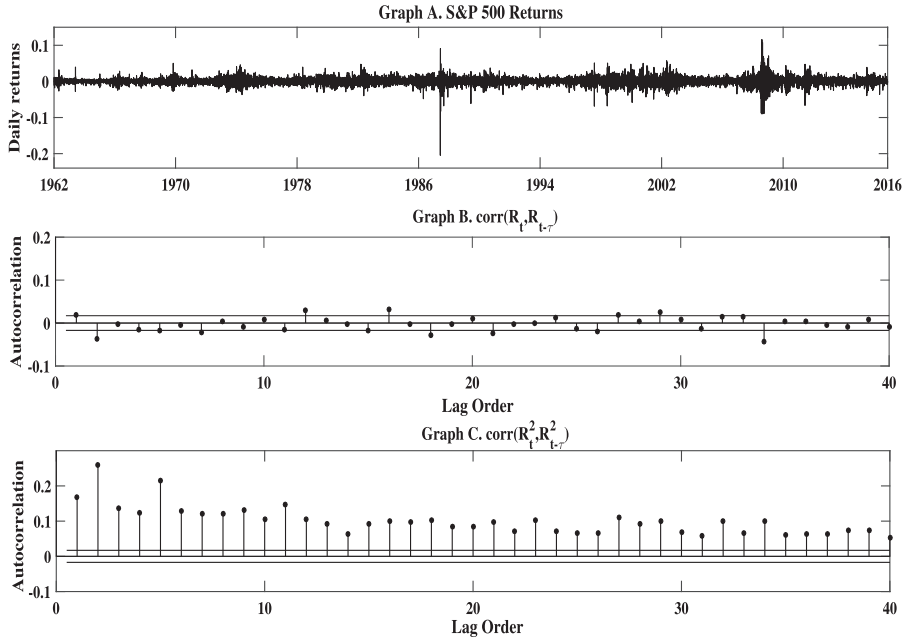

FIGURE 1 Daily S&P 500 returns and the autocorrelations of returns and squared returns

Table 2 shows the MLEs of the parameters for the IG and HN GARCH models as well as their standard errors under the physical measure ($\boldsymbol{P}$).

The properties of the HN and IG GARCH models are reported at the bottom of Table 2. The average annualized volatility is approximately 0.17. The variance persistence is about 0.960, which is in line with the existing results. The leverage coefficient for the HN GARCH model is $-9.955\text{E}-04$ and that for the IG GARCH model is $-1.100\text{E}-03$. The unconditional variance $E[\bar{h}]$ is $8.444\text{E}-05$ for the HN GARCH model and $8.669\text{E}-05$ for the IG GARCH model. The estimate of η is statistically significant and negative, which implies that the IG GARCH model has captured the negative skewness in the S&P 500 returns distribution. All these results confirm the commonly found stylized facts (Cont, 2001) such as volatility clustering, heavy tails, and the leverage effect.

In terms of the above properties in the S&P 500 returns, the HN GARCH model is quite similar to the IG GARCH model. However, their log-likelihood values are fairly different. The log-likelihood value of the IG GARCH model is about 70 points larger than that of the HN GARCH model. This confirms that the IG GARCH model can provide a much better fit than the HN GARCH model. The results of SBC and AIC information criteria show that the IG GARCH model is preferred. Consistent with the literature, a likelihood ratio test also supports the IG GARCH model, at a significance level of 1%. These are reported in Table 2.

Moreover, due to possible overfitting in the IG model, we also report the results of 1-day, 1-week, and 1-month ahead volatility forecasts for the period from January 4, 2010 to December 30, 2016 in Figure 2 and Table 3. We forecast the last 500 observations using the rolling window scheme[2] and $\sqrt{r_t^2}$ as a proxy for the real volatility. According to the results, we can see that the IG GARCH model has better predictive power for 1-day ahead and 1-week ahead volatility relative to the HN GARCH model.

Figure 3 plots the autocorrelation of model residuals and squared model residuals. Graphs A and B in Figure 3 show the autocorrelation residuals z_t and squared residuals z_t^2 of the HN GARCH model. Graphs C and D show the autocorrelation of residuals $\varepsilon_t = \frac{y_t - E(y_t)}{\text{Std}(y_t)}$ and squared residuals ε_t^2 of the IG GARCH model. Comparing these graphs, we find that the IG GARCH model captures the persistency in squared returns much better than the HN GARCH model does. In contrast, the HN GARCH model shows large autocorrelations in squared residuals z_t^2, not conforming to i.i.d. assumptions.

We report the dynamic conditional moments of the S&P 500 returns under the IG GARCH model in Figure 4, using the estimated parameters in Table 2. Graph A shows the annualized conditional return volatility for the IG GARCH model, whose dynamic path is quite similar to that of the HN GARCH model. Thus, to save space, we have not

[2]These results are based on the suggestion of an anonymous referee.

TABLE 2 Estimation of daily S&P 500 returns (1962–2016)

	HN GARCH			IG GARCH	
Parameters	**Estimate**	**Standard errors**	**Parameters**	**Estimate**	**Standard errors**
λ	1.966E + 00	2.366E − 05	ζ	1.637E + 03	5.510E − 02
ω	1.999E − 09	6.658E − 14	w	1.641E − 10	9.943E − 12
β	8.841E − 01	6.638E − 07	b	−1.931E + 01	4.700E − 03
α	3.157E − 06	4.763E − 11	a	2.472E + 07	7.713E + 03
γ	1.477E + 02	1.500E − 04	c	4.125E − 06	6.292E − 10
			η	−6.118E − 04	2.068E − 08
$E[\bar{h}]$	8.444E − 05		$E[\bar{h}]$	8.669E − 05	
Properties			Properties		
Avg. volatility	0.1670		*Avg. volatility*	0.1686	
Persistence	0.9626		*Persistence*	0.9601	
Coef. leverage	−9.955E − 04		*Coef. leverage*	−1.100E − 03	
Loglikelihoods	46,162.3		*Loglikelihoods*	46,230.4	
SBC	−92,277		*SBC*	−92,404	
AIC	−92,314		*AIC*	−92,449	
LR-test p-value			LR-test p-value	0.000***	

Note. This table shows parameter estimates, standard errors, and properties of the HN and IG GARCH models. The sample consists of daily S&P 500 returns for the period from July 3, 1962 to December 30, 2016. SBC $= k \ln N - 2 \ln L$ and AIC $= 2k - 2 \ln L$, where k is the number of parameters, N is the number of observations, and $\ln L$ is the log-likelihood.

***Significance at the 1% level.

displayed the conditional return volatility for the HN GARCH model. Graph B shows the conditional return skewness, which is quite close to the unconditional skewness reported in Table 1. Compared with the unconditional excess return kurtosis of 20.693 in Table 1, the conditional return kurtosis from the IG GARCH model, displayed in Graph C, cannot precisely describe the dynamic path of the excess return kurtosis. We may further incorporate the jump diffusion in the return process.[3]

4 | VIX VALUATION THEORY

Because the stock price can jump to an infinite set of values in a single period, the EMM is not unique in the discrete-time framework. In this section, we establish the existence of a risk-neutral measure, in which all of an asset's returns are equal to the risk-free interest rate. First, we propose a pricing kernel that provides more general valuation results than the existing literature. Then, we compute the EMM coefficient and derive the risk-neutral dynamics of the return process. Unlike the risk-neutralization approach adopted of Christoffersen et al. (2006), we do not first attempt to characterize the relationship between physical density and risk-neutral density. Fortunately, this pricing kernel also allows the return process under the risk-neutral measure ($\boldsymbol{Q}$) to be similar to that under the physical measure ($\boldsymbol{P}$).

4.1 | The pricing kernel

Following the affine asset pricing literature (Bates, 2006; Christoffersen et al., 2012), we assume that pricing kernel π_{t+1} follows the affine dynamic

[3]See, for example, Jump-GARCH in Christoffersen, Jacobs, and Ornthanalai (2012) and Ornthanalai (2014).

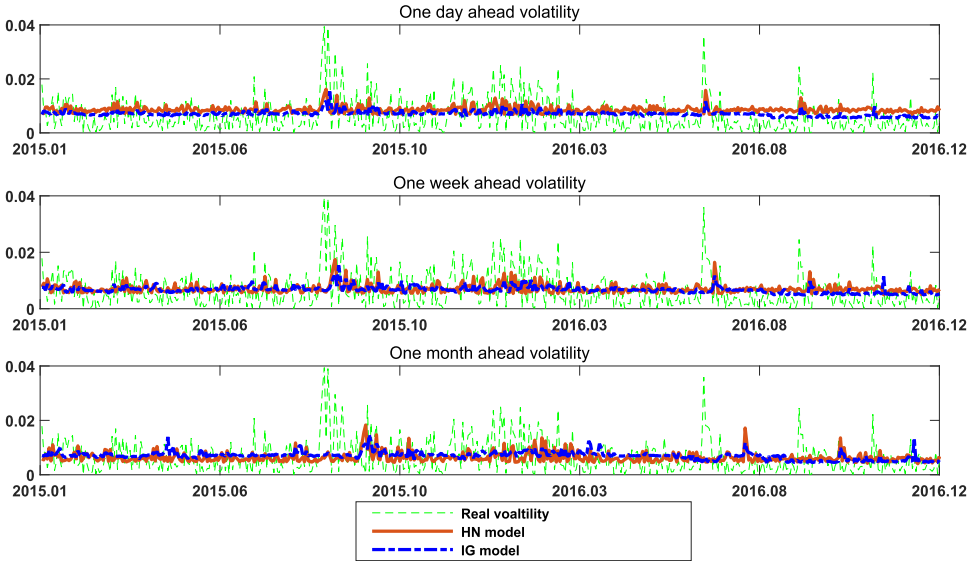

FIGURE 2 1-Day, 1-week, 1-month ahead volatility forecasts for HN and IG GARCH models, $\sqrt{r_t^2}$ proxy for the real volatility [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 3 Mean absolute errors of volatility forecasts (January 2015 to December 2016)

	One day ahead	One week ahead	One month ahead
HN GARCH (0%)	5.461	4.955	4.852
IG GARCH (0%)	4.816	4.809	4.982

Note. This table shows mean absolute error (*MAE*) between predicated volatility and real volatility ($\sqrt{r_t^2}$). The data are daily S&P 500 returns for the period from January 4, 2010 to December 30, 2016. We forecast the last 500 observations using the scheme of the rolling window.

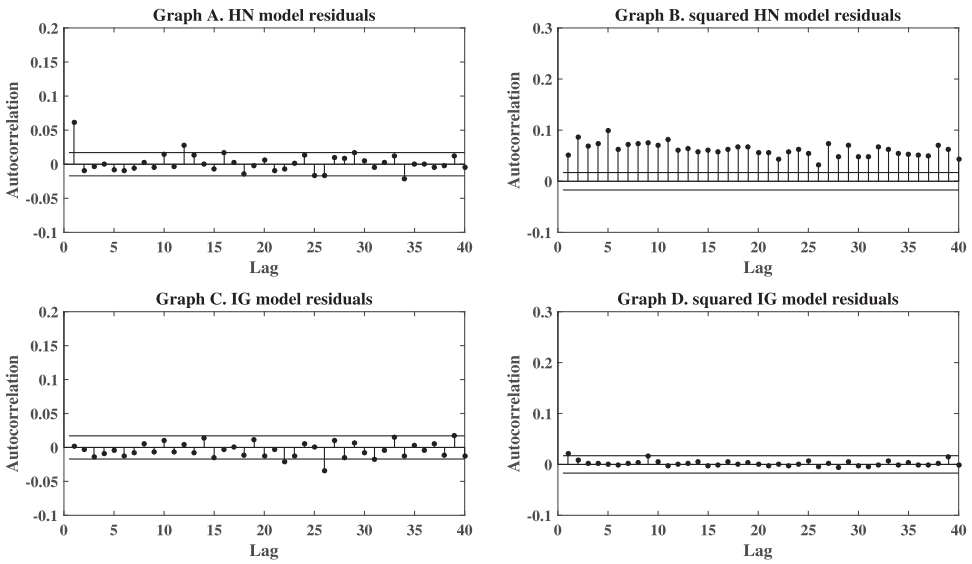

FIGURE 3 The autocorrelations of model residuals and squared model residuals for the HN GARCH model (z_t and z_t^2) and the IG GARCH model ($\varepsilon_t = \frac{y_t - E(y_t)}{\text{Std}(y_t)}$ and ε_t^2). The horizontal lines denote ±2 standard error bands. The parameter estimates are from Table 2

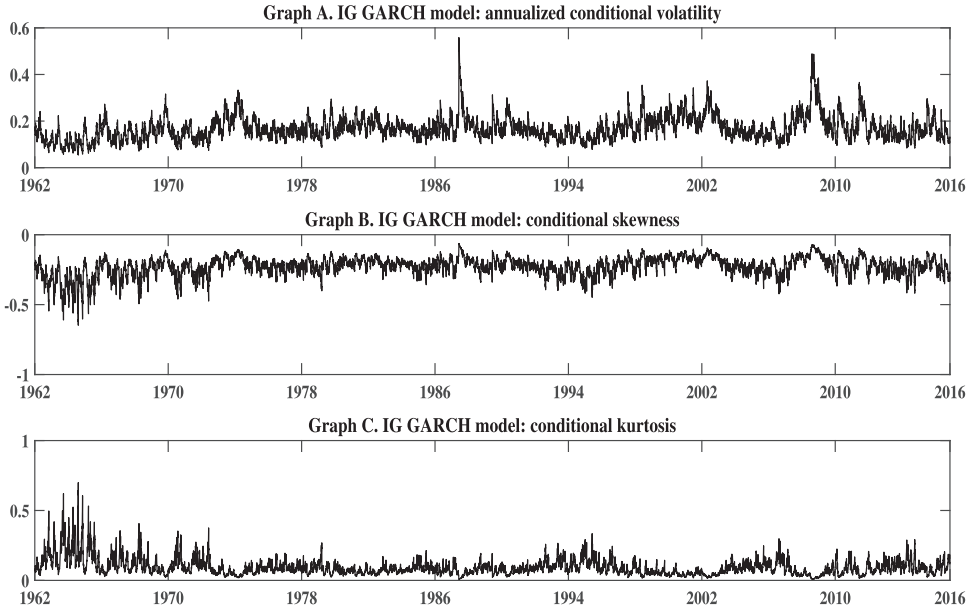

FIGURE 4 The conditional moments of S&P 500 returns. The parameter estimates are from Table 2

$$\log\left(\frac{\pi_{t+1}}{\pi_t}\right) = \alpha_{t+1} - \gamma R_{t+1}, \tag{15}$$

where R_{t+1} is the log stock return and γ is linked to the diffusive risk premium.

By substituting the dynamics of R_{t+1} into Equation (15), we get following expression:

$$\log\left(\frac{\pi_{t+1}}{\pi_t}\right) = \alpha_{t+1} - \gamma r - \gamma\zeta h_{t+1} - \gamma\eta y_{t+1}, \tag{16}$$

where the coefficient α_t is determined so that $E_{t-1}[\pi_t/\pi_{t-1}]$ equals the risk-free rate. Then, we have

$$\frac{\pi_{t+1}}{\pi_t} = \frac{\exp(r - \gamma\eta y_{t+1})}{E_t[\exp(-\gamma\eta y_{t+1})]}. \tag{17}$$

4.2 | EMM

Motivated by the pricing kernel in Equation (17), we specify the conditional Radon–Nikodym derivative as follows:

$$Z_{t+1} = \frac{\exp(\Lambda\eta y_{t+1})}{E_t[\exp(\Lambda\eta y_{t+1})]}, \tag{18}$$

where ηy_{t+1} is the inverse Gaussian component in Equation (1). Λ is defined as the EMM coefficient that captures the wedge between the $\boldsymbol{P}$ measure and the $\boldsymbol{Q}$ measure.

To make the $\boldsymbol{Q}$ measure an EMM, the following requirement is needed.

Proposition 4.1 *If the return process under the $\boldsymbol{P}$ measure is given by Equation (1), the $\boldsymbol{Q}$ measure defined by the Radon–Nikodym derivative Z_{t+1} in Equation (18) is an EMM if and only if*

$$\frac{\sqrt{1 - 2\Lambda\eta} - \sqrt{1 - 2(\Lambda\eta + \eta)}}{\eta^2} + \zeta = 0 \tag{19}$$

Proof See Appendix A. □

Note that the EMM coefficient Λ can be solved analytically by Equation (19)

$$\Lambda = -\frac{1}{2\eta}\left(\frac{(\eta^3\zeta^2+2)^2}{4\eta^2\zeta^2}-1\right) \quad \text{with} \quad \eta<0 \tag{20}$$

Therefore, the estimates of the physical parameters η and ζ can determine the EMM coefficient Λ that establishes the connection between the $\boldsymbol{P}$ measure and the $\boldsymbol{Q}$ measure.

4.3 | Risk-neutral process

After characterizing an EMM, we can derive the corresponding $\boldsymbol{Q}$ measure for the stochastic components of returns using a change of measure.

Proposition 4.2 *If the stochastic process is ηy_{t+1} under the physical measure $\boldsymbol{P}$, the stochastic process under risk-neutral measure $\boldsymbol{Q}$ can be written as $\eta^* y^*_{t+1}$ according to the Radon–Nikodym derivative Z_{t+1} in Equation (18), where*

$$y_{t+1} \sim \mathrm{IG}\left(\frac{h_{t+1}}{\eta^2}\right) \text{ and } y^*_{t+1} \sim \left(\frac{h^*_{t+1}}{\eta^{*2}}\right) \tag{21}$$

with

$$\eta^* = \frac{\eta}{1-2\Lambda\eta} \text{ and } h^*_{t+1} = \frac{h_{t+1}}{(1-2\Lambda\eta)^{\frac{3}{2}}}$$

Proof See Appendix B. □

After a change of measure from $\boldsymbol{P}$ to $\boldsymbol{Q}$, we find that, generally, the risk-neutral scale parameter and variance will be different from the physical scale parameter and variance. Now, we are ready to derive the risk-neutral process required for the VIX futures valuation.

Proposition 4.3 *Under the risk-neutral measure $\boldsymbol{Q}$, the return follows the process*

$$R_{t+1} \equiv \log\left(\frac{S_{t+1}}{S_t}\right) = r + \zeta^* h^*_{t+1} + \eta^* y^*_{t+1} \tag{22}$$

and the variance dynamic is

$$h^*_{t+1} = w^* + b h^*_t + c^* y^*_t + \frac{a^* h^{*2}_t}{y^*_t} \tag{23}$$

where

$$h_{t+1} = \Pi h^*_{t+1}, \quad \zeta^* = \frac{\sqrt{1-2\eta^*}-1}{\eta^{*2}}, \quad w^* = \frac{w}{\Pi},$$
$$c^* = \frac{c\eta^*}{\Pi\eta}, \quad a^* = \frac{a\Pi\eta}{\eta^*}, \quad \Pi = (1-2\Lambda\eta)^{\frac{3}{2}}, \quad \eta y_{t+1} = \eta^* y^*_{t+1}$$

Proof See Appendix C. □

Note that the discounted price process in Equation (22) is a martingale where $\zeta^* h^*_{t+1}$ is the compensating term for inverse Gaussian components. Ultimately, we confirm that the risk-neutral process of the IG GARCH process takes the same form as the physical process.

4.4 | The implied VIX index

According to the CBOE white paper, The VIX is defined as

$$\left(\frac{\text{VIX}_t}{100}\right)^2 = \frac{2}{\tau}\sum_i \frac{\Delta K_i}{K_i^2} e^{r\tau} Q(K_i) - \frac{1}{\tau}\left[\frac{F}{K_0} - 1\right]^2, \tag{24}$$

where $\tau = 22/252$, r is the risk-free interest rate, Q_i is the midpoint of the bid-ask spread for each option with strike K_i, and K_0 is the first strike below the forward index level, F. Following Lin (2007) in Appendix A, this formula can be converted to

$$\left(\frac{\text{VIX}_t}{100}\right)^2 = -\frac{2}{\tau}E^Q\left[\log\left(\frac{S_T}{F}\right)\right] = 2r - \frac{2}{\tau}E_t^Q\left[\log\left(\frac{S_T}{S_t}\right)\right], \tag{25}$$

where $\tau = T - t$ and $F = S_t e^{r\tau}$. Substituting the return dynamic in Equation (22), we have

$$\left(\frac{\text{VIX}_t}{100}\right)^2 = \frac{2}{\tau}(-\zeta^* - 1/\eta^*)\sum_{k=1}^{n} E_t^Q[h^*_{t+k}] \tag{26}$$

where h^*_{t+k} is daily variance of the S&P 500 returns under the $\boldsymbol{Q}$ measure.

The VIX approximates the 30-day variance swap rate on the S&P 500 Index (Bollerslev, Gibson, & Zhou, 2011; Carr & Wu, 2006; Kanniainen, Lin, & Yang, 2014).[4] Therefore, it can be interpreted as measuring the risk-neutral expectation of integrated variance within a month in the absence of jumps and in discrete time,

$$\begin{aligned}\left(\tfrac{\text{VIX}_t}{100}\right)^2 &\cong \tfrac{1}{n}\sum_{k=1}^{n} E_t^Q[h^*_{t+k}] \times 252 \\ &= \text{Vix}_t(n) \times 252\end{aligned} \tag{27}$$

where $\text{Vix}_t(n) = \frac{1}{n}\sum_{k=1}^{n} E_t^Q[h^*_{t+k}]$ is defined as a proxy for VIX_t in terms of daily variance. To convert the daily variance into the annualized variance, we set the annualizing factor as 252. The n generally equals 22, treated as 22 trading days for a month.

We can compute the risk-neutral conditional mean of futures variance in Equation (27) using the risk-neutral variance dynamics in Equation (23). The IG GARCH model provides the following relationship between the implied $\text{Vix}_t(n)$ and conditional variance.

Proposition 4.4 *If the S & P 500 returns follow the IG GARCH process presented in Proposition 4.3 under the* $\boldsymbol{Q}$ *measure, then the implied* Vix_t *at time t is a weighted average of the unconditional variance* $\bar{h}^*$ *and next period variance* h^*_{t+1},

$$\text{Vix}_t(n) = \bar{h}^*\left(1 - \frac{1-\Gamma^n}{n(1-\Gamma)}\right) + h^*_{t+1}\frac{1-\Gamma^n}{n(1-\Gamma)} \tag{28}$$

[4]According to the results $\eta = -6.118e-04$ and $\zeta = 1.637e+03$ in Table 2, we can have $2(-\zeta^* - 1/\eta^*) = 1$. Because the final results are not sensitive to this consideration, we fix the $2(-\zeta^* - 1/\eta^*)$ at one, noting that complete results are available from the authors upon request.

where

$$\bar{h}^* = \frac{w^* + \eta^{*4} a^*}{1 - \left(a^* \eta^{*2} + b + \frac{c^*}{\eta^{*2}}\right)}$$

Γ *is a measure of variance persistence under the* $\boldsymbol{Q}$ *measure*

$$\Gamma = a^* \eta^{*2} + b + \frac{c^*}{\eta^{*2}}$$

Therefore, the corresponding implied VIX_t *is given by*

$$\mathrm{VIX}_t = 100\sqrt{\tilde{a} + \tilde{b} h^*_{t+1}} \tag{29}$$

with

$$\tilde{a} = 252 \times \left(1 - \frac{1 - \Gamma^n}{n(1 - \Gamma)}\right)\bar{h}^* \text{ and } \tilde{b} = 252 \times \frac{1 - \Gamma^n}{n(1 - \Gamma)}$$

Proof See Appendix D. □

If the S&P 500 returns follow the HN GARCH dynamics, the risk-neutral process is

$$R_{t+1} \equiv \log \frac{S_{t+1}}{S_t} = r - \frac{1}{2} h^*_{t+1} + \sqrt{h^*_{t+1}} z^*_{t+1} \tag{30}$$

$$h^*_{t+1} = \omega + \beta h^*_t + \alpha\left(z^*_{t+1} - \gamma^* \sqrt{h^*_t}\right) \tag{31}$$

$$\text{with } z^*_{t+1} \sim N(0, 1) \quad \gamma^* = \gamma + \lambda + 0.5 \tag{32}$$

Therefore, the corresponding $\bar{h}^*$ and Γ of the HN GARCH model will be $\tilde{b}^* = \frac{(\omega + \alpha)}{(1 - (\beta + \alpha\gamma^{*2}))}$ and $\Gamma = \beta + \alpha\gamma^{*2}$. The final expression of VIX_t under the HN GARCH model is in line with the result presented by Wang et al. (2017).

4.5 | The pricing formula of VIX futures

The price of a VIX futures contract at the current time t can be computed as its conditional expectation at maturity time T under the risk-neutral measure. According to Proposition 4.4, we derive the pricing formula of VIX futures $F_{t,T}$, where $T - t$ is the time to maturity

$$F(t, T) = E^Q_t[\mathrm{VIX}_T | \mathcal{F}_t] = E^Q_t[100\sqrt{\tilde{a} + \tilde{b} h^*_{T+1}} | \mathcal{F}_t] \tag{33}$$

where $\mathcal{F}$ is the filtration. Following Zhu and Lian (2012), we can change the VIX pricing formula to

$$\begin{aligned} F(t, T) &= \frac{100}{2\sqrt{\pi}} \int_0^\infty \frac{1 - E^Q_t\left[e^{-s\left(\tilde{a} + \tilde{b} h^*_{T+1}\right)}\right]}{s^{3/2}} ds \\ &= \frac{100}{2\sqrt{\pi}} \int_0^\infty \frac{1 - e^{-s\tilde{a}} E^Q_t[e^{-s\tilde{b} h^*_{T+1}}]}{s^{3/2}} ds \end{aligned} \tag{34}$$

The term $E^Q[e^{-s\tilde{b}h^*_{T+1}}]$ is the moment-generating function of h^*_{T+1} under the $\boldsymbol{Q}$ measure. Fortunately, we can solve this function analytically by guessing the form of the moment-generating function.

Proposition 4.5 *If the conditional variance h^*_{t+m} (m days from now) at time t follows the dynamics of the IG GARCH model, then the moment-generating function of h^*_{t+m} is exponentially affine in h^*_{t+1}, and the expression is*

$$f(\phi, m, h^*_{t+1}) = E^Q_t[e^{\phi h_{t+m+1}}] = e^{A(\phi,m)+B(\phi,m)h^*_{t+1}}. \tag{35}$$

The function $A(\phi, m)$ and $B(\phi, m)$ are given by the following iterative relationship,

$$A(\phi, m+1) = A(\phi, m) + A_J(B(\phi, m))$$
$$B(\phi, m+1) = B_J(B(\phi, m))$$

where

$$A_J(B(\phi, m)) = w^*B(\phi, m) - \tfrac{1}{2}\ln(1 - 2a^*\eta^{*4}B(\phi, m))$$
$$B_J(B(\phi, m)) = \tfrac{1}{\eta^{*2}} + bB(\phi, m) - \sqrt{\left(\tfrac{1}{\eta^{*4}} - 2a^*B(\phi, m)\right)(1 - 2c^*B(\phi, m))}$$

with the following initial conditions

$$A(\phi, 0) = 0 \text{ and } B(\phi, 0) = \phi$$

Proof See Appendix E. □

Under the HN GARCH model, the corresponding $A_J(B(\phi, m))$ and $B_J(B(\phi, m))$ can be computed as

$$A_J(B(\phi, m)) = -\tfrac{1}{2}\ln(1 - 2\alpha B(\phi, m)) + \omega B(\phi, m)$$
$$B_J(B(\phi, m)) = \tfrac{\alpha\gamma^{*2}B(\phi, m)}{1 - 2\alpha B(\phi, m)} + \beta B(\phi, m)$$

Therefore, the VIX futures price under the IG GARCH model and the HN GARCH model can be rewritten as

$$F(t, T) = \frac{100}{2\sqrt{\pi}}\int_0^\infty \frac{1 - e^{-s\tilde{a}}f\left(-s\tilde{b}, T - t, \frac{(\mathrm{VIX}_t/100)^2 - \tilde{a}}{\tilde{b}}\right)}{s^{\frac{3}{2}}}\,ds \tag{36}$$

Equation (36) develops the pricing relationship between the contemporaneous VIX index and the VIX futures. It is also worth noting that the parameters $\tilde{a}$ and $\tilde{b}$ and the moment-generating function $f(\cdot)$ correspond to the IG and HN GARCH models, respectively. They are different for different models.

5 | PERFORMANCE OF FUTURES VALUATION

We now discuss the VIX futures valuation of the IG GARCH model and the benchmark HN GARCH model. We first introduce the futures data used in our empirical research and then estimate the parameters using the MLE method. Finally, we analyze the in- and out-of-sample pricing performance of VIX futures.

5.1 | Futures data

In this section, we describe the VIX futures data, which are split into in-sample data and out-of-sample data. The in-sample data consist of contracts from 2004 to 2016, and the out-of-sample data consist of contracts from 2017 that have expired. All of the VIX futures contracts are directly collected from CBOE.

Following Bakshi, Cao, and Chen (1997) and Zhu and Lian (2012), we apply several filters to the raw data. We eliminate futures with maturities less than 5 days, open interest less than 200 points, and prices less than 0.5, considering the tick size of 0.01. Guided by the CBOE Futures Exchange rescaling method, we rescale the VIX futures prices for the period from March 26, 2004 to March 25, 2007, by dividing the prices by 10. Finally, we obtain 21,902 in-sample observations and 2,207 out-of-sample observations.

Table 4 shows descriptive statistics for VIX futures prices sorted by maturity, the bias between VIX index and VIX futures prices, and the VIX level.

Panel A of Table 4 sorts the data by maturity. The mean of VIX futures prices is roughly steady across different maturities. Panel B sorts the data by bias between the VIX index and the VIX futures price. Most futures contracts are concentrated in bias between −5 and 0. Panel C sorts the data by VIX index level. The fluctuation and price of VIX futures increase as the VIX index rises.

5.2 | Fitting futures prices

In this section, we estimate the parameters of the HN and IG GARHCH models under the risk-neutral measure by directly maximizing the futures pricing error log-likelihoods. The corresponding log-likelihood function is

$$L_{\text{futures}} = -\frac{N}{2}\ln(2\pi\sigma^2) - \frac{1}{2}\sum_{j=1}^{N}\frac{\left(F_j^{\text{MKT}} - F_j^{\text{MOD}}\right)^2}{\sigma^2}, \tag{37}$$

where N is the total number of futures observations, F_j^{MKT} and F_j^{MOD} denote the VIX futures market price and the VIX futures theoretical price calculated by the HN and IG GARCH models, and σ^2 represents the variance of futures pricing errors. Recall that the VIX futures theoretical price is calculated by Equation (36) with the market VIX index.

Table 5 shows the results of the futures-based estimation. Note that we first estimate the unconditional risk-neutral variance $E^Q[\bar{h}]$ and then imply the ω of the HN GARCH model and the w of the IG GARCH model.

We report the mean absolute error (*MAE*) and the root mean square error (*RMSE*) at the bottom of Table 5. The *MAE* of the IG GARCH model is 3.0327, which is an improvement of 2.84% over the benchmark HN GARCH model. The *RMSE* of the IG GARCH model is 3.7539, less than the 3.8313 *RMSE* of the HN GARCH model. Furthermore, we find that the variance components under the risk-neutral measure in Table 5 tend to be more persistent than those under the physical measure in Table 2. In contrast, the absolute values of the leverage coefficients in Table 5 are smaller than those in Table 2. These findings are in line with the results in the literature.

In summary, the estimated results from the models confirm that the IG GARCH model combining conditional skewness of the inverse Gaussian distribution offers a better VIX futures pricing method than the HN GARCH model does.

Table 6 further shows the in-sample futures price fit in *RMSE* across maturity, bias, and VIX level, with the parameters reported in Table 5.

Panel A of Table 6 shows the *RMSE* of the models by maturity. The futures pricing performance of the HN GARCH model is better than that of the IG GARCH model when the DTM is less than 50 days. In contrast, the IG GARCH model outperforms the HN GARCH model for longer DTM, especially for medium- and long-term futures contracts.

Looking across the columns of Panel B of Table 6, we find that the IG GARCH model fits the futures better than the HN GARCH model for BIAS $\leq$ 5, which covers almost 96% of the total samples. Fitting the futures is relatively difficult for both models when the bias is larger than 5.

Panel C in Table 6 shows the *RMSE* by the VIX index level. During periods of extreme VIX levels (too low or too high), the IG GARCH model performs better. Note that the difficulty of futures valuation increases at high VIX levels, which is also the case in the options valuation literature.

TABLE 4 In-sample VIX futures contracts for the period 2004–2016

Panel A. By maturity	**DTM ≤ 50**	**50 < DTM ≤ 80**	**80 < DTM ≤ 120**	**120 < DTM**	**All**
Observations	6,989	3,906	5,008	5,999	21,902
Mean	20.48	22.06	22.70	21.82	21.64
Maximum	66.23	54.67	47.07	45	66.23
Minimum	10.37	12.53	13.38	13.52	10.37
Standard deviation	7.80	6.85	6.27	5.20	6.70
Panel B. By bias	**BIAS ≤ −5**	**−5 < BIAS ≤ 0**	**0 < BIAS ≤ 5**	**5 < BIAS**	All
Observations	4,441	13,577	2,951	933	21,902
Mean	23.64	19.51	24.45	34.20	21.64
Maximum	35.15	54.46	63.88	66.23	66.23
Minimum	14.98	10.37	10.99	17.04	10.37
Standard deviation	4.32	5.35	8.00	8.94	6.70
Panel C. By VIX level	**VIX ≤ 15**	**15 < VIX ≤ 20**	**20 < VIX ≤ 30**	**30 < VIX**	All
Observations	7,997	6,702	5,154	2,049	21,902
Mean	16.36	20.68	25.50	35.65	21.64
Maximum	27.65	31.00	34.55	66.23	66.23
Minimum	10.37	13.50	16.71	20.07	10.37
Standard deviation	2.31	3.23	3.41	7.11	6.70

Note. This table shows the summary of statistics of the VIX futures contracts from 2004 to 2016. BIAS = VIX index − VIX futures price. DTM denotes the days to maturity.

TABLE 5 Estimation on VIX futures contracts (2004–2016)

	HN GARCH			**IG GARCH**	
Parameters	**Estimate**	**Standard errors**	**Parameters**	**Estimate**	**Standard errors**
ω	2.878E − 06		w^*	2.170E − 06	
β	9.910E − 01	2.289E − 04	b	7.208E − 01	1.100E − 03
α	1.001E − 07	1.277E − 09	a^*	2.544E + 04	1.041E + 03
γ^*	5.438E + 00	1.995E − 02	c^*	5.674E − 07	8.246E − 09
			η^*	−1.700E − 03	6.819E − 06
$E^Q[\bar{h}]$	3.310E − 04	2.488E − 06	$E^Q[\bar{h}]$	3.557E − 04	2.893E − 06
Properties			Properties		
Persistence	0.9910		*Persistence*	0.9933	
Coef. leverage	−1.088E − 06		*Coef. leverage*	−2.162E − 04	
Futures errors			Futures errors		
MAE	3.1213		*MAE*	3.0327	
RatiotoHN	1		*RatiotoHN*	0.9716	
RMSE	3.8313		*RMSE*	3.7539	
RatiotoHN	1		*RatiotoHN*	0.9798	

Note. This table shows the futures-based parameter estimates, standard errors, properties, and pricing errors of the HN and IG GARCH models. The VIX futures are for the period from 2004 to 2016.

TABLE 6 In-sample *RMSE* by maturity, bias, and VIX level for model parameters estimated on VIX futures

Panel A. *RMSE* by maturity				
Model	**DTM ≤ 50**	**50 < DTM ≤ 80**	**80 < DTM ≤ 120**	**120 < DTM**
HN GARCH	3.4283	3.6321	3.9322	4.2896
IG GARCH	3.5260	3.5772	3.7294	4.1243
Ratio	1.0285	0.9849	0.9484	0.9614
Panel B. *RMSE* by bias				
Model	**BIAS ≤ −5**	**−5 < BIAS ≤ 0**	**0 < BIAS ≤ 5**	**5 < BIAS**
HN GARCH	4.4906	3.4989	3.3752	5.8659
IG GARCH	4.4555	3.3574	2.9788	6.6497
Ratio	0.9922	0.9596	0.8826	1.1336
Panel C. *RMSE* by VIX level				
Model	**VIX ≤ 15**	**15 < VIX ≤ 20**	**20 < VIX ≤ 30**	**30 < VIX**
HN GARCH	3.2884	3.1145	4.3061	5.9541
IG GARCH	3.1339	3.2557	4.1844	5.7792
Ratio	0.9530	1.0453	0.9717	0.9706

Note. This table shows the *RMSE* futures fit across maturity, bias (VIX index − VIX futures price), and the VIX index for the models estimated in Table 5. The futures price data are collected from CBOE for the period 2004–2016. DTM denotes days to maturity. Ratio represents the ratio of the *RMSE* of the IG GARCH model to the *RMSE* of the HN GARCH model.

The principal feature of using the IG GARCH model lies in the fat-tailed innovation compared to the normal innovation underlying the HN GARCH model. Hence, the IG GARCH model may outperform the HN GARCH model in extreme periods of market volatility, illiquidity, default and funding risks, and other market news event, mainly for medium- to long-term futures contracts.

Hence, we highlight the in-sample performance results in those "joint" subsamples that lie at the intersection of bias, maturity, and VIX levels in Table 7. To determine the significance of the difference of the pricing errors, the Diebold and Mariano (2002) test is also adopted to test whether the differences between the pricing errors of the benchmark model "HN GARCH" and the test model "IG GARCH" are significant. A positive and significant value of DM statistic

TABLE 7 The in-sample performance results in the joint subsamples

		80 ≤ DTM ≤ 150			
Panel A. Model	**0 < BIAS ≤ 5**	**VIX ≤ 15**	**15 < VIX ≤ 25**	**25 < VIX ≤ 35**	**35 < VIX**
HN GARCH	3.3752	3.5588	3.5869	4.8837	6.5198
IG GARCH	2.9788	3.1587	3.5406	4.3253	6.3571
DM statistics	20.06**	69.33**	11.77**	26.70**	2.21**
Observations	2,951	2,833	4,019	983	551
		150 < DTM			
Panel B. Model	**0 < BIAS ≤ 5**	**VIX ≤ 15**	**15 < VIX ≤ 25**	**25 < VIX ≤ 35**	**35 < VIX**
HN GARCH	3.3752	5.2478	3.8214	3.2373	5.0233
IG GARCH	2.9788	5.1725	3.7746	3.1313	4.3434
DM statistics	20.06**	11.94**	5.69**	2.37**	2.19**
Observations	2,951	1,224	1,292	182	52

Note. This table shows in-sample performance results in those "joint" subsamples that lie at the intersection of bias, maturity, and VIX levels. The futures price data are collected from CBOE from 2004 to 2016. DTM denotes the days to maturity. BIAS = VIX index − VIX futures price.

**DM statistic that is significant at the 5% significance level.

TABLE 8 Estimation on VIX futures contracts on the extreme days of the 2008 financial crisis

	HN GARCH			IG GARCH	
Parameters	**Estimate**	**Standard errors**	**Parameters**	**Estimate**	**Standard errors**
ω	1.302E − 05		w^*	7.1048E − 06	
β	9.650E − 01	1.700E − 03	b	7.235E − 01	1.200E − 03
α	1.060E − 07	2.098E − 09	a^*	2.544E + 04	1.249E + 03
γ^*	5.3524E + 00	1.358E − 01	c^*	2.495E − 07	2.520E − 09
			η^*	−3.000E − 03	2.239E − 06
$E^Q[\bar{h}]$	3.749E − 04	9.195E − 06	$E^Q[\bar{h}]$	3.458E − 04	9.419E − 06
Properties			Properties		
Persistence	0.9650		*Persistence*	0.9739	
Coef. leverage	−1.1350E − 06		*Coef. leverage*	5.6999E − 04	
Futures errors			Futures errors		
MAE	2.7595		*MAE*	2.3429	
RatiotoHN	1		*RatiotoHN*	0.8490	
RMSE	3.4851		*RMSE*	3.1903	
RatiotoHN	1		*RatiotoHN*	0.9154	

Note. This table shows the futures-based parameter estimates, standard errors, properties, and pricing errors of the HN and IG GARCH models. The VIX futures are for the period from September 1, 2008 to October 30, 2008.

implies that the HN GARCH model has larger pricing errors than the IG GARCH model. And ** represents a DM statistic that is significant at the 5% significance level.

When BIAS is between 0 and 5, the *RMSE* of the IG GARCH model is 2.9788, which is an improvement of 11.74% over the HN GARCH model. The DM statistic is 20.06**, suggesting that the IG GARCH model significantly outperforms the HN GARCH model. The *RMSE* metric and the test statistic under column "80 ≤ DTM ≤ 150" of Panel A reveals that the IG GARCH model is superior to the HN GARCH model in all four VIX categories. Also, it shows that the IG GARCH model performs significantly better than the HN GARCH model in each of the four VIX categories under column "150 < DTM" of Panel B, that is, the test statistics are positive and significant.

5.3 | Pricing performance during the 2008 financial crisis

The 2008 financial crisis hit its peak in September and October 2008. Several major institutions either failed, were acquired under duress, or were subjected to government takeover. These included Lehman Brothers, Merrill Lynch, Fannie Mae, Freddie Mac, Washington Mutual, Wachovia, Citigroup, and AIG. We re-estimate these pricing models based on the VIX futures from September 1, 2008 to October 30, 2008, which yields 307 observations. Table 8 shows the results of the futures-based estimation during the extreme period of the 2008 financial crisis.

At the bottom of Table 8, we report the *MAE* and *RMSE*. The *MAE* for the IG GARCH model is 2.3429, compared with 2.7595 for the benchmark HN GARCH model. This is an improvement of 15.10%. The *RMSE* of the IG GARCH model is 3.1903, which outperforms 3.4851 of the HN GARCH model by about 8.46%. These empirical findings indicate that incorporating the conditional skewness into the GARCH model does substantially improve the pricing accuracy of VIX futures on the extreme days of the 2008 financial crisis.

We also dissect the overall *RMSE* results reported in Table 9 by sorting data by maturity, bias, and VIX index level. Looking across columns of Table 9, we see that the IG GARCH model has lower *RMSE* than the HN GARCH model across categories, suggesting that the benefits offered by the IG GARCH model are not restricted to any particular subset during the extreme period of the 2008 financial crisis.

Table 10 shows the pricing performance results in those "joint" subsamples that lie at the intersection of bias, maturity, and VIX levels during the 2008 financial crisis. Again, the IG GARCH model significantly outperforms the HN GARCH model in each of the joint subsamples. Compared with the results reported in Table 7, the improvement of pricing performance between the IG GARCH model and the HN GARCH model on the extreme days of the 2008

TABLE 9 *RMSE* by maturity, bias, and VIX level for model parameters estimated on futures contracts on the extreme days of the 2008 financial crisis

Panel A. *RMSE* by maturity				
Model	**DTM ≤ 50**	**50 < DTM ≤ 80**	**80 < DTM ≤ 120**	**120 < DTM**
HN GARCH	2.6371	3.7680	3.8751	3.7884
IG GARCH	2.6022	3.6698	3.3668	3.3794
Observations	98	66	87	56
Panel B. *RMSE* by bias				
Model	**BIAS ≤ −5**	**−5 < BIAS ≤ 0**	**0 < BIAS ≤ 5**	**5 < BIAS**
HN GARCH	–	2.7102	2.3622	3.7282
IG GARCH	–	1.5660	1.4557	3.5548
Observations	–	39	34	234
Panel C. *RMSE* by VIX level				
Model	**VIX ≤ 15**	**15 < VIX ≤ 20**	**20 < VIX ≤ 30**	**30 < VIX**
HN GARCH	–	–	2.5444	3.7218
IG GARCH	–	–	1.4911	3.5456
Observations	–	–	71	236

Note. This table shows the *RMSE* futures fit across maturity, bias (VIX index − VIX futures price), and VIX index for the models estimated in Table 8. The futures price data are collected from CBOE from September 1, 2008 to October 30, 2008. DTM denotes the days to maturity.

financial crisis has been largely increased, suggesting that the fat-tailed innovation underlying the IG GARCH model plays an important role in the VIX futures pricing in the extreme periods.

5.4 | The out-of-sample results

In this section, we analyze the out-of-sample performance of the models using the expiring VIX futures contracts from 2017. Table 11 displays the basic statistical features of these 2,207 observations across maturity, bias, and the VIX index level.

TABLE 10 The pricing performance results in the joint subsamples during the 2008 financial crisis

		80 ≤ DTM ≤ 150			
Panel A. Model	**0 < BIAS ≤ 5**	**VIX ≤ 15**	**15 < VIX ≤ 25**	**25 < VIX ≤ 35**	**35 < VIX**
HN GARCH	2.3622	–	2.9297	2.2612	4.8125
IG GARCH	1.4557	–	1.5660	1.1780	4.2256
DM statistics	7.32**	–	28.56**	19.33**	4.98**
Observations	34	–	21	26	72
		150 < DTM			
Panel B. Model	**0 < BIAS ≤ 5**	**VIX ≤ 15**	**15 < VIX ≤ 25**	**25 < VIX ≤ 35**	**35 < VIX**
HN GARCH	2.3622	–	3.4422	2.7095	1.1249
IG GARCH	1.4557	–	2.2068	1.5547	0.7684
DM statistics	7.32**	–	17.71**	12.16**	1.25
Observations	34	–	7	9	10

Note. This table shows pricing performance results in those "joint" subsamples that lie at the intersection of bias, maturity, and VIX levels on the extreme days of the 2008 financial crisis. The futures price data are collected from September 1, 2008 to October 30, 2008. DTM denotes the days to maturity. BIAS = VIX index − VIX futures price.

**DM statistic that is significant at the 5% significance level.

Looking at Table 11, we find that the out-of-sample futures prices across maturity, bias, and the VIX level tend to be more centralized than the in-sample futures prices in Table 4, with the data showing a smaller maximum, a larger minimum, and a smaller standard deviation. No observations are in the range: $5 \leq$ BIAS and $30 \leq$ VIX. Not surprisingly, the market price volatility tends to be low after the 2008 financial crisis.

Table 12 shows the out-of-sample performance of the HN and IG GARCH models by maturity, bias, and VIX level.

Panel A of Table 12 shows the out-of-sample *RMSE* across DTM. We see an almost 19% improvement for the IG GARCH model relative to the HN GARCH model in forecasting the medium- and long-term futures contracts, but the IG GARCH model performs poorly for DTM $\leq$ 50, which is in line with the in-sample findings in Table 6.

Panel B of Table 12, which displays the *RMSE* across bias, shows that the forecasting performance of the IG GARCH model is notably better than that of the HN GARCH model for each subsample. In total, the IG GARCH model has futures-pricing errors that are 8% below those of the HN GARCH model.

Panel C of Table 12 also confirms that the IG GARCH model has a better performance for out-of-sample forecasting. Note that when $20 \leq$ VIX, only 23 observations exist. As such a small sample is not representative, we do not report the corresponding *RMSE*.

Table 13 shows the out-of-sample performance results in those "joint" subsamples that lie at the intersection of bias, maturity, and VIX levels. When $0 <$ BIAS ≤ 5, the pairwise test statistic significantly favors the IG GARCH model over the HN GARCH model, that is, the test statistic is 2.31**. Similarly, the positive and significant DM statistics under columns "$80 \leq$ DTM ≤ 150" and "$150 <$ DTM" suggest that the IG GARCH model should be a better modeling choice in terms of pricing medium- and long-term VIX futures contracts.

TABLE 11 Out-of-sample VIX futures contracts from 2017

Panel A. By maturity					
	DTM $\leq$ 50	**50 < DTM $\leq$ 80**	**80 < DTM $\leq$ 120**	**120 < DTM**	**All**
Observations	787	360	480	580	2,207
Mean	13.341	15.638	17.349	19.171	16.120
Maximum	19.775	19.925	20.725	23.225	23.225
Minimum	10.375	13.075	14.5	15.675	10.375
Standard deviation	1.471	1.680	1.582	1.405	2.804
Panel B. By bias					
	BIAS ≤ -5	**-5 < BIAS $\leq$ 0**	**0 < BIAS $\leq$ 5**	**5 < BIAS**	**All**
Observations	766	1,356	85	–	2,207
Mean	18.493	14.759	16.430	–	16.120
Maximum	21.850	22.650	23.225	–	23.225
Minimum	14.375	10.375	12.225	–	10.375
Standard deviation	1.510	2.449	2.767	–	2.804
Panel C. By VIX level					
	VIX $\leq$ 15	**15 < VIX $\leq$ 20**	**20 < VIX $\leq$ 30**	**30 < VIX**	**All**
Observations	1,984	200	23	–	2,207
Mean	15.805	18.679	21.003	–	16.120
Maximum	21.175	21.875	23.225	–	23.225
Minimum	10.375	13.875	19.725	–	10.375
Standard deviation	2.683	2.221	1.1396	–	2.804

Note. This table shows the summary of statistics of VIX futures contracts from 2017. BIAS = VIX index – VIX futures price. DTM denotes the days to maturity.

TABLE 12 Out-of-sample *RMSE* by maturity, bias, and VIX level for model parameters estimated on futures contracts

Panel A. *RMSE* by maturity					
Model	**DTM ≤ 50**	**50 < DTM ≤ 80**	**80 < DTM ≤ 120**	**120 < DTM**	**All**
HN GARCH	2.0976	2.0494	2.6833	2.7386	2.4083
IG GARCH	2.3238	1.1491	2.1447	2.4494	2.2135
Ratio	1.1078	0.7071	0.7993	0.8944	0.9191
Panel B. *RMSE* by bias					
Model	**BIAS ≤ −5**	**− 5 < BIAS ≤ 0**	**0 < BIAS ≤ 5**	**5 < BIAS**	**All**
HN GARCH	2.4290	2.4494	1.8439	–	2.4083
IG GARCH	2.0248	2.2361	1.8166	–	2.2135
Ratio	0.8336	0.9130	0.9852	–	0.9191
Panel C. *RMSE* by VIX level					
Model	**VIX ≤ 15**	**15 < VIX ≤ 20**	**20 < VIX ≤ 30**	**30 < VIX**	**All**
HN GARCH	2.4495	2.0833	–	–	2.4083
IG GARCH	2.2360	1.9493	–	–	2.2135
Ratio	0.9128	0.9357	–	–	0.9191

Note. This table shows the *RMSE* futures fit across maturity, bias (VIX index − VIX futures price), and VIX index for the models estimated in Table 5. The futures price data are collected from CBOE from 2017. DTM denotes the days to maturity. Ratio represents the ratio of the *RMSE* of the IG GARCH model to the *RMSE* of the HN GARCH model.

6 | JOINT ESTIMATION ON RETURNS AND FUTURES

To better reconcile the information from the $\boldsymbol{P}$ and $\boldsymbol{Q}$ measures, the jointly fitting method has been popular in current derivative pricing; see, for example, Eraker (2004), Santa-Clara and Yan (2010), and Christoffersen et al. (2014). Combining the two likelihoods in Equations (14) and (37), we derive the total log-likelihood

$$L_{\text{joint}} = L_{\text{returns}} + L_{\text{futures}} \tag{38}$$

TABLE 13 The out-of-sample performance results in the joint subsamples

		80 ≤ DTM ≤ 150			
Panel A. Model	**0 < BIAS ≤ 5**	**VIX ≤ 15**	**15 < VIX ≤ 25**	**25 < VIX ≤ 35**	**35 < VIX**
HN GARCH	1.8439	2.8228	2.3013	–	–
IG GARCH	1.8166	2.3373	2.1548	–	–
DM statistics	2.31**	43.03**	6.73**	–	–
Observations	85	728	114	–	–
		150 < DTM			
Panel B. Model	**0 < BIAS ≤ 5**	**VIX ≤ 15**	**15 < VIX ≤ 25**	**25 < VIX ≤ 35**	**35 < VIX**
HN GARCH	1.8439	2.6070	2.2890	–	–
IG GARCH	1.8166	2.3604	2.1818	–	–
DM statistics	2.31**	26.04**	3.47**	–	–
Observations	85	185	45	–	–

Note. This table shows the out-of-sample performance results in those "joint" subsamples that lie at the intersection of bias, maturity, and VIX levels. The futures price data are collected from CBOE from 2017. DTM denotes the days to maturity. BIAS = VIX index − VIX futures price.

**DM statistic that is significant at the 5% significance level.

Recall that parameters in the physical measure can analytically obtain all the parameters needed to recover the risk-neutral dynamics, according to Section 4. We use the EMM coefficient Λ to capture the wedge between the $\boldsymbol{P}$ and $\boldsymbol{Q}$ measures. The Λ of the IG GARCH model is the expression in Equation (20), and the corresponding Λ of the HN GARCH model is $\lambda + 0.5$. Table 14 shows these results from the joint estimation on returns and futures.

The main results of Table 14 are as follows:

First, the average annualized volatility of the S&P 500 returns is higher than the average annualized volatility reported in Table 2. We find a similarly high average annualized volatility in Chorro and Fanirisoa (2016) when they jointly fit the S&P 500 returns and the VIX.

Second, the likelihood of futures favors the IG GARCH model. However, jointly fitting the IG GARCH model results in a relatively poor performance. The persistence of volatility is close to the futures-based results in Table 5. The leverage coefficients of the models are larger than those in Tables 2 and 5.

TABLE 14 Joint estimation on S&P 500 returns and VIX futures contracts

	HN GARCH			IG GARCH	
Parameters	**Estimate**	**Standard errors**	**Parameters**	**Estimate**	**Standard errors**
λ	6.191E − 01	8.440E − 02	ζ	6.562E + 01	2.341E − 07
ω	3.299E − 06	7.791E − 08	w	1.976E − 15	2.845E − 13
β	9.890E − 01	3.397E − 04	b	9.869E − 01	2.416E − 10
α	1.003E − 10	8.625E − 10	a	5.885E + 00	6.151E − 09
γ	1.075E + 02	7.862E + 00	c	1.023E − 14	6.429E − 13
γ^*	1.087E + 02		η	−2.690E − 02	2.395E − 10
			w^*	1.921E − 15	
			a^*	6.168E + 00	
			c^*	9.762E − 15	
			a^*	6.168E + 00	
			η^*	−2.640E − 02	
Λ	6.691E − 01		Λ	3.529E − 01	
$E[\bar{h}]$	2.994E − 04		$E[\bar{h}]$	3.474E − 04	
Properties			Properties		
Ave. volatility	0.3309		*Ave. volatility*	0.3121	
Persistence	0.9890		*Persistence*	0.9911	
Coef. leverage	−2.156E − 08		*Coef. leverage*	1.147E − 04	
Total log-likelihood	−3.121E + 04		Total log-likelihood	−3.745E + 04	
Futures log-likelihood	−7.191E + 04		Futures log-likelihood	−7.015E + 04	
Returns log-likelihood	4.070E + 04		Returns log-likelihood	3.270E + 04	
Futures errors			Futures errors		
MAE	3.2069		*MAE*	3.1511	
RatiotoHN	1		*RatiotoHN*	0.9826	
RMSE	3.9492		*RMSE*	3.8462	
RatiotoHN	1		*RatiotoHN*	0.9739	

Note. We jointly fit the S&P 500 returns from 1962 to 2016 and the VIX futures contracts from 2004 to 2016 and report the parameters, standard errors, properties, and futures pricing errors for the models.

Third, the joint *MAE* and *RMSE* of the IG GARCH model are smaller than those of the HN GARCH model, indicating that the IG GARCH model capturing the negative conditional skewness of the returns distribution outperforms the HN GARCH model when we price VIX futures.

Overall, the results in Table 14 again confirm that it is important to combine conditional skewness, conditional heteroskedasticity, and a leverage effect in a model, if we wish to develop a reliable valuation formula for VIX futures.

7 | CONCLUSION AND FUTURE WORK

This paper presents an analytical valuation formula for VIX futures based on the discrete-time inverse Gaussian GARCH model that adequately captures the stylized facts of conditional skewness, conditional heteroskedasticity, and a leverage effect in asset returns. Meanwhile, instead of directly assuming the form of state-price density, we develop a pricing kernel using the risk-neutralization framework of Christoffersen et al. (2010). This framework allows us to derive the risk-neutral process without fully characterizing the economic environment.

Some important conclusions are obtained from the empirical analysis on the S&P 500 returns and the VIX futures prices. First, compared with the HN GARCH model with normal innovations, the IG GARCH model with nonnormal innovations is more suitable for describing the return dynamics. Second, although its in-sample pricing performance is mixed, IG GARCH model outperforms the HN GARCH model as a whole. The IG GARCH model substantially reduces the pricing errors during the 2008 financial crisis, relative to the benchmark model. Third, an out-of-sample test significantly favors the IG GARCH model over the HN GARCH model, particularly for futures contracts with a longer maturity. The IG model's pricing errors for medium- and long-term VIX futures contracts are 11–29% lower than the errors obtained by the HN model. Meanwhile, the IG model slightly outperforms HN model when contracts display in small bias and high VIX level.

This study can be extended in several ways. First, motivated by the relatively poor performance under high VIX levels, our work can be extended to incorporate the Lévy jumps developed in Ornthanalai (2014). Second, according to the research by Meddahi and Renault (2004) and Hao and Zhang (2013), VIX can be calculated under a broad class of GARCH models that combine other stylized empirical facts in financial markets. Third, considering that macrofinance variables have effects on conditional volatility, it would be useful to model VIX with the MacroGARCH proposed by Dorion (2016). Finally, applying the models to the realized variance would be interesting. We leave these topics for further research.

ACKNOWLEDGMENTS

We are grateful to Bob Webb (editor), an anonymous referee, Qingma Dong, Zhiyuan Pan, Yuhuang Shang, Guangwei Zhu, and Ji Chen whose comments substantially improved the paper. This study was supported by a grant from the National Natural Science Foundation of China (71473200). Xinglin Yang and Peng Wang are both at the Institute of Chinese Financial Studies, Southwestern University of Finance and Economics, Chengdu, Sichuan, China.

ORCID

Xinglin Yang http://orcid.org/0000-0002-0511-6219

REFERENCES

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *The Journal of Finance*, *52*, 2003–2049.

Bates, D. S. (2006). Maximum likelihood estimation of latent affine processes. *The Review of Financial Studies*, *19*, 909–965.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*, 307–327.

Bollerslev, T., Gibson, M., & Zhou, H. (2011). Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities. *Journal of Econometrics*, *160*, 235–245.

Carr, P., & Wu, L. (2006). A tale of two indices. *The Journal of Derivatives*, *13*, 13–29.

Chorro, C. & Fanirisoa Rahantamialisoa, H. (2016). Option valuation with IG_GARCH model and an U-shaped pricing kernel. Retrieved from https://halshs.archives-ouvertes.fr/halshs-01400242/

Christoffersen, P., Elkamhi, R., Feunou, B., & Jacobs, K. (2010). Option valuation with conditional heteroskedasticity and nonnormality. *The Review of Financial Studies*, *23*, 2139–2183.
Christoffersen, P., Feunou, B., Jacobs, K., & Meddahi, N. (2014). The economic value of realized volatility: Using high-frequency returns for option valuation. *Journal of Financial and Quantitative Analysis*, *49*, 663–697.
Christoffersen, P., Heston, S., & Jacobs, K. (2006). Option valuation with conditional skewness. *Journal of Econometrics*, *131*, 253–284.
Christoffersen, P., Jacobs, K., & Ornthanalai, C. (2012). Dynamic jump intensities and risk premiums: Evidence from S&P500 returns and options. *Journal of Financial Economics*, *106*, 447–472.
Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, *1*, 223–236.
Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *20*, 134–144.
Dorion, C. (2016). Option valuation with macro-finance variables. *Journal of Financial and Quantitative Analysis*, *51*, 1359–1389.
Dotsis, G., Psychoyios, D., & Skiadopoulos, G. (2007). An empirical comparison of continuous-time models of implied volatility indices. *Journal of Banking & Finance*, *31*, 3584–3603.
Duan, J.-C. (1995). The GARCH option pricing model. *Mathematical Finance*, *5*, 13–32.
Eraker, B. (2004). Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *The Journal of Finance*, *59*, 1367–1403.
Exchange, C. B. O. (2015). The CBOE volatility index–VIX. Retrieved from https://www.cboe.com/micro/vix/vixwhite.pdf
Folks, J. L., & Chhikara, R. S. (1978). The inverse Gaussian distribution and its statistical application: A review. *Journal of the Royal Statistical Society. Series B (Methodological)*, *40*, 263–289.
Grünbichler, A., & Longstaff, F. A. (1996). Valuing futures and options on volatility. *Journal of Banking & Finance*, *20*, 985–1001.
Hao, J., & Zhang, J. E. (2013). GARCH option pricing models, the CBOE VIX, and variance risk premium. *Journal of Financial Econometrics*, *11*, 556–580.
Heston, S. L., & Nandi, S. (2000). A closed-form GARCH option valuation model. *The Review of Financial Studies*, *13*, 585–625.
Kanniainen, J., Lin, B., & Yang, H. (2014). Estimating and using GARCH models with VIX data for option valuation. *Journal of Banking & Finance*, *43*, 200–211.
Lin, Y.-N. (2007). Pricing VIX futures: Evidence from integrated physical and risk-neutral probability measures. *Journal of Futures Markets*, *27*, 1175–1217.
Meddahi, N., & Renault, E. (2004). Temporal aggregation of volatility models. *Journal of Econometrics*, *119*, 355–379.
Mencía, J., & Sentana, E. (2013). Valuation of VIX derivatives. *Journal of Financial Economics*, *108*, 367–391.
Ornthanalai, C. (2014). Levy jump risk: Evidence from options and returns. *Journal of Financial Economics*, *112*, 69–90.
Patton, A. J. (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics*, *2*, 130–168.
Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, *47*, 527–556.
Patton, A. J., & Timmermann, A. (2007). Properties of optimal forecasts under asymmetric loss and nonlinearity. *Journal of Econometrics*, *140*, 884–918.
Santa-Clara, P., & Yan, S. (2010). Crashes, volatility, and the equity premium: Lessons from S&P 500 options. *The Review of Economics and Statistics*, *92*, 435–451.
Wang, T., Shen, Y., Jiang, Y., & Huang, Z. (2017). Pricing the CBOE VIX futures with the Heston-Nandi GARCH model. *Journal of Futures Markets*, *37*, 641–659.
Zhang, J. E., & Zhu, Y. (2006). VIX futures. *Journal of Futures Markets*, *26*, 521–531.
Zhu, S.-P., & Lian, G.-H. (2012). An analytical formula for VIX futures and its applications. *Journal of Futures Markets*, *32*, 166–190.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Yang X, Wang P. VIX futures pricing with conditional skewness. *J Futures Markets*. 2018;1–26. https://doi.org/10.1002/fut.21925

APPENDIX A

For an equivalent martingale measure (EMM), we need to impose that the expected return S_{t+1}/S_t must equal the risk-free interest rate

$$E_t^Q\left[\frac{S_{t+1}}{S_t}\right] = E_t\left[Z_{t+1}\frac{S_{t+1}}{S_t}\right] = \exp(r) \tag{A.1}$$

Then, we combine the Radon–Nikodym derivative Z_{t+1} in Equation (18) and the return process in Equation (1)

$$\frac{\exp(r + \zeta h_{t+1}) E_t[\exp(\Lambda\eta + \eta) y_{t+1}]}{E_t[\exp(\Lambda\eta y_{t+1})]} = \exp(r) \tag{A.2}$$

and take logs

$$\log \frac{E_t[\exp((\Lambda\eta + \eta) y_{t+1})]}{E_t[\exp(\Lambda\eta y_{t+1})]} + \zeta h_{t+1} = 0. \tag{A.3}$$

To solve the conditional equation above, we need to use the moment-generating function (MGF) of the inverse Gaussian component ηy_{t+1} in Equation (1). From the MGF of an inverse Gaussian variable, we can obtain the MGF of ηy_{t+1}

$$E_t[\exp(\phi\eta y_{t+1})] = \exp\left(\frac{(1 - \sqrt{1 - 2\phi\eta}) h_{t+1}}{\eta^2}\right), \tag{A.4}$$

substituting

$$\frac{E_t[\exp((\Lambda\eta + \eta) y_{t+1})]}{E_t[\exp(\Lambda\eta y_{t+1})]} = \frac{\exp\left(\frac{h_{t+1}}{\eta^2}(1 - \sqrt{1 - 2(\Lambda\eta + \eta)})\right)}{\exp\left(\frac{h_{t+1}}{\eta^2}(1 - \sqrt{1 - 2\Lambda\eta})\right)} \tag{A.5}$$

in the EMM restriction (A.3). Simplifying and collecting terms, we obtain

$$\left(\frac{\sqrt{1 - 2\Lambda\eta} - \sqrt{1 - 2(\Lambda\eta + \eta)}}{\eta^2} + \zeta\right) h_{t+1} = 0 \tag{A.6}$$

This can be solved by equating the coefficient of h_{t+1} to zero, which will give the conclusion of Equation (19).

APPENDIX B

We will prove this by retrieving the form of MGF under the $\boldsymbol{Q}$ measure through the Radon–Nikodym derivative Z_{t+1} in Equation (18). First, the MGF of ηy_{t+1} under the risk-neutral probability measure is specified as $E_t^Q[\exp(\phi\eta y_{t+1})]$. After a simple change of measure, it becomes

$$\begin{aligned} E_t^Q[\exp(\phi\eta y_{t+1})] &= E_t[Z_{t+1} \exp(\phi\eta y_{t+1})] = E_t\left[\frac{\exp(\Lambda\eta y_{t+1})}{E_t[\exp(\Lambda\eta y_{t+1})]} \exp(\phi\eta y_{t+1})\right] \\ &= E_t\left[\frac{\exp((\Lambda\eta + \phi\eta) y_{t+1})}{E_t[\exp(\Lambda\eta y_{t+1})]}\right]. \end{aligned} \tag{B.1}$$

Using the MGF of ηy_{t+1} in Equation (A.4), we can simplify Equation (B.1)

$$E_t^Q[\exp(\phi\eta y_{t+1})] = \exp\left[\frac{h_{t+1}}{\eta^2}(1-\sqrt{1-2(\Lambda+\phi)\eta}) - \frac{h_{t+1}}{\eta^2}(1-2\sqrt{1-2\Lambda\eta})\right]$$
$$= \exp\left[\frac{h_{t+1}}{(1-2\Lambda\eta)^{\frac{3}{2}}}\frac{(1-2\Lambda\eta)^2}{\eta^2}\left(1-\sqrt{1-\frac{2\eta}{1-2\Lambda\eta}\phi}\right)\right]$$
$$= \exp\left[\frac{h_{t+1}^*}{\eta^{*2}}(1-\sqrt{1-2\eta^*\phi})\right]. \tag{B.2}$$

where

$$\eta^* = \frac{\eta}{1-2\Lambda\eta} \quad \text{and} \quad h_{t+1}^* = \frac{h_{t+1}}{(1-2\Lambda\eta)^{\frac{3}{2}}}.$$

We denote the inverse Gaussian component in the $\boldsymbol{Q}$ measure by $y_{t+1}^* \sim IG(\frac{h_{t+1}^*}{\eta^{*2}})$. Accordingly, we have proved that the $E_t^Q[\exp(\phi\eta y_{t+1})]$ is the moment-generating function of the stochastic process $\eta^* y_{t+1}^*$.

APPENDIX C

Using the conclusion of Proposition 4.2, the return process in Equation (1) under the $\boldsymbol{Q}$ measure can be written as

$$R_{t+1} \equiv \log\left(\frac{S_{t+1}}{S_t}\right) = r + \zeta h_{t+1} + \eta^* y_{t+1}^*. \tag{C.1}$$

The risk-neutral variance dynamic

$$h_{t+1} = w + bh_t + c\frac{\eta^*}{\eta}y_t^* + a\frac{\eta}{\eta^*}\frac{h_t^{*2}}{y_t^*}. \tag{C.2}$$

Note that the shock of risk-neutral distribution is $y_{t+1}^* \sim \text{IG}(h_{t+1}^*/\eta^{*2})$. Furthermore, because the $\boldsymbol{Q}$ measure is constructed such that the discounted process of S_t in Equation (C.1) is martingale, we need to have

$$\zeta h_{t+1} = -\frac{h_{t+1}^*}{\eta^{*2}}(1-\sqrt{1-2\eta^*}) = \zeta^* h_{t+1}^*, \tag{C.3}$$

where $\zeta^* = \frac{\sqrt{1-2\eta^*}-1}{\eta^{*2}}$ and h_{t+1}^* is the conditional variance under the risk-neutral measure. The return process in Equation (C.1) simplifies to

$$R_{t+1} \equiv \log\left(\frac{S_{t+1}}{S_t}\right) = r + \zeta^* h_{t+1}^* + \eta^* y_{t+1}^*. \tag{C.4}$$

Furthermore, we reparameterize the variance dynamics in Equation (C.2) through the following transformations:

$$h_{t+1} = \Pi h_{t+1}^*, \quad w^* = \frac{w}{\Pi}, \quad c^* = \frac{c\eta^*}{\Pi\eta}, \quad a^* = \frac{a\Pi\eta}{\eta^*}.$$

when $\Pi = (1-2\Lambda\eta)^{\frac{3}{2}}$. The variance dynamics become

$$h_{t+1}^* = w^* + bh_t^* + c^* y_t^* + \frac{a^* h_t^{*2}}{y_t^*}. \tag{C.5}$$

APPENDIX D

For $k \geq 1$, given the available information at time t, the conditional expectation of variance under the risk-neutral probability measure is

$$E_t^Q(h_{t+k}^*) = w^* + \eta^{*4}a^* + \left(a^*\eta^{*2} + b + \frac{c^*}{\eta^{*2}}\right)E_t^Q(h_{t+k-1}^*) = \Delta + \Gamma E_t^Q(h_{t+k-1}^*). \tag{D.1}$$

$$\begin{aligned} E_t^Q(h_{t+k-1}^*) &= \Delta + \Gamma E_t^Q(h_{t+k-2}^*) \\ &\vdots \\ E_t^Q(h_{t+2}^*) &= \Delta + \Gamma h_{t+1}^* \end{aligned}$$

with

$$\Delta = w^* + \eta^{*4}a^* \quad \text{and} \quad \Gamma = a^*\eta^{*2} + b + \frac{c^*}{\eta^{*2}}.$$

Combining these equations in Equation (D.1), we obtain

$$E_t^Q(h_{t+k}^*) = \Delta\frac{1-\Gamma^{k-1}}{1-\Gamma} + h_{t+1}^*\Gamma^{k-1} \tag{D.2}$$

By the definition of $\text{Vix}_t(n)$, we have

$$\begin{aligned} \text{Vix}_t(n) &= \frac{1}{n}\sum_{k=1}^{n} E_t^Q(h_{t+k}^*) \\ &= \frac{1}{n}\left[\Delta\sum_{k=1}^{n}\frac{1-\Gamma^{k-1}}{1-\Gamma} + \sum_{k=1}^{n}\Gamma^{k-1}h_{t+1}^*\right] \\ &= \frac{\Delta}{1-\Gamma}\left(1 - \frac{1-\Gamma^n}{n(1-\Gamma)}\right) + h_{t+1}^*\frac{1-\Gamma^n}{n(1-\Gamma)} \\ &= \bar{h}^*\left(1 - \frac{1-\Gamma^n}{n(1-\Gamma)}\right) + h_{t+1}^*\frac{1-\Gamma^n}{n(1-\Gamma)} \end{aligned} \tag{D.3}$$

where $\bar{h}^* = \frac{\Delta}{1-\Gamma} = \frac{w^* + \eta^{*4}a^*}{1-\left(a^*\eta^{*2} + b + \frac{c^*}{\eta^{*2}}\right)}$. Accordingly, the $\text{Vix}_t(n)$ is the weight average of unconditional variance $\bar{h}^*$ and conditional variance h_{t+1}^* under the $\boldsymbol{Q}$ measure.

APPENDIX E

Suppose that the form of the moment-generating function is

$$E_t^Q\left[e^{\phi h_{t+m+2}^*}\right] = e^{A(\phi,m+1)+B(\phi,m+1)h_{t+1}^*} \tag{E.1}$$

with the initial condition $A(\phi, 0) = 0$ and $B(\phi, 0) = \phi$. Then, we also assume the following recursive relationship

$$A(\phi, m+1) = A(\phi, m) + A_J(B(\phi, m))$$
$$B(\phi, m+1) = B_J(B(\phi, m))$$

Then,

$$E_t^Q\left[e^{\phi h^*_{t+m+2}}\right] = E_t^Q\left[E_{t+1}^Q\left[e^{\phi h^*_{t+m+2}}\right]\right] = E_t^Q\left[e^{A(\phi,m)+B(\phi,m)h^*_{t+2}}\right] \tag{E.2}$$

Defining $s = B(\phi, m)$, we obtain

$$E_t^Q[e^{sh^*_{t+2}}] = E_t^Q\left[e^{s\left(w^*+bh^*_{t+1}+c^*y^*_{t+1}+\frac{a^*{h^*_{t+1}}^2}{y^*_{t+1}}\right)}\right]$$
$$= e^{sw^*+sbh^*_{t+1}}E_t^Q\left[e^{sc^*y^*_{t+1}+\frac{sa^*{h^*_{t+1}}^2}{y^*_{t+1}}}\right]$$

where $y^*_{t+1} \sim \mathrm{IG}(\frac{h^*_{t+1}}{\eta^{*2}})$. Define $\Phi = sc^*$, $\Theta = sa^*{h^*_{t+1}}^2$, and $\delta^* = \frac{h^*_{t+1}}{\eta^{*2}}$. We have

$$E_t^Q\left[e^{sc^*y^*_{t+1}+\frac{sa^*{h^*_{t+1}}^2}{y^*_{t+1}}}\right] = E_t^Q\left[e^{\Phi y^*_{t+1}+\frac{\Theta}{y^*_{t+1}}}\right]$$
$$= \int_0^\infty e^{\Phi x+\frac{\Theta}{x}}\frac{\delta^*}{\sqrt{2\pi x^3}}e^{-\frac{(\sqrt{x}-\delta^*/\sqrt{x})^2}{2}}\,dx$$
$$= \frac{\delta^*}{\sqrt{\delta^{*2}-2\Theta}}\exp(\delta^* - \sqrt{(\delta^{*2}-2\Theta)(1-2\Phi)})\int_0^\infty \frac{\sqrt{\delta^{*2}-2\Theta}}{\sqrt{2\pi x^3}}\exp\left[-\frac{(\sqrt{1-2\Phi}\cdot x-\sqrt{\delta^{*2}-2\Theta})^2}{2x}\right]dx$$
$$= \frac{\delta^*}{\sqrt{\delta^{*2}-2\Theta}}\exp(\delta^* - \sqrt{(\delta^{*2}-2\Theta)(1-2\Phi)})$$

where $\frac{\sqrt{\delta^{*2}-2\Theta}}{\sqrt{2\pi x^3}}\exp[-\frac{(\sqrt{1-2\Phi}\cdot x-\sqrt{\delta^{*2}-2\Theta})^2}{2x}]$ is a probability distribution function (pdf) of the first passage time distribution presented in Folks and Chhikara (1978). Therefore, the integral part is 1.

Collecting these results gives

$$E_t^Q\left[e^{sh^*_{t+2}}\right] = \exp\left[w^*s - \frac{1}{2}\ln(1-2sa^*\eta^{*4})\right].\exp\left[\left(\frac{1}{\eta^{*2}} + bs - \sqrt{\left(\frac{1}{\eta^{*4}} - 2a^*s\right)(1-2c^*s)}\right)h^*_{t+1}\right] \tag{E.3}$$

By comparison, we have

$$A_J(s) = w^*s - \tfrac{1}{2}\ln(1-2sa^*\eta^{*4})$$
$$B_J(s) = \tfrac{1}{\eta^{*2}} + bs - \sqrt{\left(\tfrac{1}{\eta^{*4}} - 2a^*s\right)(1-2c^*s)}$$