Do VIX Futures Contribute to the Valuation of VIX Options?

Chen Tong[a] **Tianyi Wang**[b*]

[a]*Department of Finance, School of Economics & Wang Yanan Institute for Studies in Economics (WISE), Xiamen University*

[b]*School of Banking and Finance, University of International Business and Economics*

July 15, 2021

Abstract

Basically, as the VIX index is non-tradable, most investors use the exchange-traded VIX futures to hedge their exposure in VIX options. However, the information role of VIX futures in pricing VIX options is not fully explored. In this paper, we utilize a simple discrete-time model of VIX dynamics with long-memory and asymmetric jumps to incorporate VIX futures in the pricing framework. We provided extensive empirical evidence based on CBOE VIX options from 2006 to 2020 that support the new framework's significant performance gains over existing frameworks based on SPX daily returns, realized variance, or VIX index itself. Among these models, the futures-based model provides the best pricing performance, with a reduction in pric up to 50% (compared with the VIX-based model).

Keywords: VIX Futures, VIX Options, VIX Derivatives
JEL Classification: G12, G13, C51, C52

*Corresponding author (tianyiwang@uibe.edu.cn). Tianyi Wang acknowledges financial support from the Fund of the National Natural Science Foundation of China (71871060).

1 Introduction

Introduced by the Chicago Board Options Exchange (CBOE), the Volatility Index (VIX) is the leading volatility indicator for the U.S. stock market. To meet investors demand on hedge volatility risk directly, VIX futures and VIX options are introduced by CBOE in 2004 and 2006 accordingly. Together with VIX ETFs such as VXX and VIXY, the volatility derivatives market has become a new important component of the capital market (Luo et al., 2019). The average daily trading volume of both instruments increases dramatically over the last decade and inspired a growing literature on pricing issues.

Technically, there are two major strategies for pricing VIX derivatives. The first (indirect) approach starts with a model for the underlying S&P 500 index dynamics (returns and/or realized volatility). The dynamics is then linked to VIX index through a pricing kernel. Zhang and Zhu (2006) made the first attempt to price VIX futures based on the classical continuous-time Heston model. Adding jumps and/or mean-reverting features are also investigated by Lin (2007), Sepp (2008), Zhang et al. (2010), Zhu and Lian (2012), etc. Rely on varieties of GARCH models, Wang et al. (2017), Yang and Wang (2018), Yang et al. (2019), Wang and Wang (2021), Cao et al. (2020), Tong and Huang (2021) also use this indirect approach to price VIX derivatives. The second (direct) approach starts with VIX index dynamics and skips the risk neutralization process as well. This approach includes Grünbichler and Longstaff (1996), Goard and Mazur (2013), Park (2016), Kaeck and Alexander (2013), Mencía and Sentana (2013), Psychoyios et al. (2010), Cao et al. (2020) and Jing et al. (2020) and others. For these models, the current level of VIX always serves as a state variable in the VIX derivatives pricing formulas.

In this paper, we focus on the providing a simple and effective method to price VIX options. To do this, we follow the second approach as we build our model directly on the logarithm of VIX index using ARMA(p,q) process with jumps in innovations. The jump part is assumed to follow the double exponential distribution (Kou, 2002) to model the possible asymmetry and the explicit pricing formula using VIX index is then provided through the Fourier inverse transformation. As documented in Yin et al. (2021) using the special case of the current model, the implied VIX futures price is an exponential affine function of VIX index. Using similar technique, we provide a explicit link between VIX options and futures that enable us to further use VIX futures as state variables for pricing VIX options.

Instead of introducing additional parameters or volatility factors to boost the pricing performance as previous researches did, switching state variables from VIX series to VIX futures panel greatly increases the number of states and, therefore, the pricing accuracy without complicating the model. We provided extensive empirical evidence based on CBOE VIX options from 2006 to 2020 that support the new framework's significant performance gains over existing frameworks based on SPX daily returns, realized variance, or VIX index itself. Among these models, the futures-based model provides the best pricing performance, with a reduction in RMSE up to 50% (compared with the VIX-based model).

Our paper contributes to the literature on pricing VIX options using discrete-time models from several aspects. Firstly, it provides a simple but rich framework for VIX option pricing with long-memory and mean-reverting in log features. The framework includes special cases such as heterogeneous autoregressive (HAR) model (Yin et al., 2021) and the continuous time limit is linked to log-OU process (Detemple and Osakwe, 2000). Secondly, we not only provide the conventional VIX based pricing formula but also provide a VIX futures based pricing formula. The latter formula is motivated by the fact that investor use tradable VIX futures to hedge VIX options rather than the non-tradable VIX index. Our method is related to Lin (2013) which use forward VIX (calculated from VIX futures and VIX term structure) to price VIX options with two notable differences: 1) our model is much simpler as we do not need VIX term structure as input; 2) our model builds on logarithm of VIX which which has received considerable empirical support in modeling VIX. Thirdly, our research provide a way to quantitatively demonstrate the richer information embedded in VIX futures. Results suggest VIX futures should be used as a default choice of state variables.

The rest of the paper is organized as follows. Section 2 introduces the model setup and derives the pricing formula for VIX options; Section 3 describes the data and discusses the model estimation results; Section 4 presents the empirical results; and Section 5 concludes.

2 The model

Our model starts with the logarithm of VIX (i.e., $V_t = \log(\mathrm{VIX}_t)$) which is a common technique to deal with nonnegative series (Nelson, 1991), restore normality, and reduce the adverse effect of extreme values (Andersen et al., 2003). It is also a popular form to model VIX in continuous-time models such as Detemple and Osakwe (2000), Mencía and Sentana (2013)and Park (2016). The dynamics of V_t in risk-neutral measure $\mathbb{Q}$ is assumed to follow an ARMA(p,q) process with jumps:

$$V_t = \mu + \sum_{i=1}^{p} \beta_i V_{t-i} + \sum_{j=1}^{q} \alpha_i \varepsilon_{t-j} + \varepsilon_t$$

where the innovation ε_t is the mixture of a standard normal distribution z_t and a double exponential jump J_t:

$$\varepsilon_t \equiv \delta z_t + J_t$$

The double exponential jump J_t is introduced by Kou (2002) to describe the asymmetry of positive and negative jumps between period $t-1$ and t:

$$J_t = \sum_{j=0}^{N_t} e_{j,t}, \quad e_{j,t} \sim \mathrm{Kou}(p_u, p_d, \eta_u, \eta_d), \quad N_t \sim Poisson(\lambda)$$

where λ is the common intensity of jumps. Each jump is a mixture of two exponential distributions whose average jump size equals to $1/\eta_1$ and $1/\eta_2$ respectively. p_u and p_d represent the probability of upward and downward jumps. J_t is widely used in pricing options such as in Kou and Wang (2004), Yang (2018) and others. The Kou's specification can be reduced to no jumps $J_t = 0$ and asymmetric jumps $\eta_1 = \eta_2$, both of them can be tested with real data.

Our framework has several cases as well. For example, if we omit moving average part and constrain autoregressive parameters as follows:

$$\beta_1 = \beta_d \quad \beta_i = \frac{1}{4}\beta_w \ \ (i = 2, \dots, 5) \quad \beta_i = \frac{1}{17}\beta_m \ \ (i = 6, \dots, 22)$$

The ARMA(p,q) will reduce to a HAR model described in Yin et al. (2021). Together with two additional quarterly and yearly components:

$$\beta_i = \frac{1}{41}\beta_q \ \ (i = 23, \dots, 63) \quad \beta_i = \frac{1}{189}\beta_y \ \ (i = 64, \dots, 252)$$

The ARMA(p,q) will reduce to a similar HARG-Y model introduced in Huang et al. (2019). We also test the model with short memory settings where we only keep the first order lag of V_t and ε_t.

3 The pricing strategies

By definition, the theoretical value of VIX call option at time t with a maturity date at time T equals:

$$C_t = e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}\left[\max\left(\text{VIX}_T - K, 0\right)\right]$$

In this section, we provide two pricing strategies using VIX index and VIX futures as state variables accordingly. To begin with, the moment generating function (MGF) of V_t under the ARMA(p,q) with jumps follows an exponential affine structure on lagged (log) volatility index $V_{t-1}, \dots, V_{t+1-p}$ and innovations $\varepsilon_{t-1}, \dots, \varepsilon_{t+1-q}$ with parameters calculated through simple iteration. We summarized this in the following proposition.

PROPOSITION 1. *Under the proposed model, the moment generating function of V_{t+m} at time t follows the exponential affine structure in $V_{t-1}, \dots, V_{t+1-p}$ as well as $\varepsilon_{t-1}, \dots, \varepsilon_{t+1-q}$.*

$$g(t,k,s) = \mathbb{E}_t^{\mathbb{Q}}\left(\exp\left(sV_{t+k}\right)\right) = \exp\left(A(k,s) + \sum_{i=1}^{p} B_i(k,s)V_{t+1-i} + \sum_{j=1}^{q} C_j(k,s)\varepsilon_{t+1-j}\right)$$

where $k = T - t$ and $A(k,s)$, $B_i(k,s)$, $C_j(k,s)$ are defined by the following recursive formulas:

$$\begin{aligned}
A(k+1,s) &= A(k,s) + B_1(k,s)\mu + \Psi(B_1(k,s) + C_1(k,s)) \\
B_i(k+1,s) &= \begin{cases} B_{i+1}(k,s) + B_1(k,s)\beta_i & 1 \le i < p \\ B_1(k,s)\beta_i & i = p \end{cases} \\
C_i(k+1,s) &= \begin{cases} C_{i+1}(k,s) + B_1(k,s)\alpha_i & 1 \le i < q \\ B_1(k,s)\alpha_i & i = q \end{cases}
\end{aligned}$$

Where $\Psi(s) \equiv \log \mathbb{E}_t^{\mathbb{Q}}(\exp(s\varepsilon_{t+1}))$. *Initial conditions are*

$$A(0,s) = 0, \quad B_i(0,s) = \begin{cases} s & i = 1 \\ 0 & 1 < i < p \end{cases}, \quad C_j(0,s) = 0 \ \ (j = 1, ..., q)$$

Proof. See the Appendix. □

The theoretical price of VIX futures at time t with a maturity T is equal to the risk-neutral expectation of VIX_T conditional on time t:

$$F_{t,T} = \mathbb{E}_t^{\mathbb{Q}}(\text{VIX}_T) = \mathbb{E}_t^{\mathbb{Q}}(\exp(V_T)) = g(t, T-t, 1)$$

Therefore, the corresponding VIX futures price implied by the model is the special case where $s = 1$.

REMARK 1. *The VIX futures price at time t and under the given model can be expressed as:*

$$F_{t,T} = \mathbb{E}_t^{\mathbb{Q}}(\exp(V_T)) = \exp\left(A(k) + \sum_{i=1}^{p} B_i(k) V_{t+1-i} + \sum_{i=1}^{q} C_i(k)\varepsilon_{t+1-i}\right)$$

$$A(0) = 0, \quad C(0) = 0, \quad B_i(0) = \begin{cases} 1 & i = 1 \\ 0 & 1 < i < p \end{cases}$$

$$\begin{aligned} A(k+1) &= A(k) + B_1(k)\mu + \Psi(B_1(k) + C_1(k)) \\ B_i(k+1) &= \begin{cases} B_{i+1}(k) + B_1(k)\beta_i & 1 \le i < p \\ B_1(k)\beta_i & i = p \end{cases} \\ C_i(k+1) &= \begin{cases} C_{i+1}(k) + B_1(k)\alpha_i & 1 \le i < q \\ B_1(k)\alpha_i & i = q \end{cases} \end{aligned}$$

If we omit the MA part, the C_j terms will be dropped and the resulting expression coincides with the one listed in Yin et al. (2021). If we further omit the long memory (i.e. $p = 1$), the VIX futures price will be solely depends on the current VIX level. In this special case, pricing options based on VIX or VIX futures are identical in theory. When $p > 1$, the difference between the two approach is fundamental.

3.1 VIX-based VIX Option Pricing Formula

Note that the VIX call option price can be rewritten as:

$$\begin{aligned} C_t &= e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}\left[\max\left(\mathrm{VIX}_T - K, 0\right)\right] \\ &= e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}\left[\max\left(\exp(V_T) - K, 0\right)\right] \\ &= e^{-r(T-t)}\left[\int_{\ln K}^{\infty}\exp(x)p(x)dx - K\int_{\ln K}^{\infty}p(x)dx\right] \end{aligned}$$

where $p(x)$ is the conditional probability density function of V_T. Taking a similar mathematical technique adopted by Heston and Nandi (2000, Proposition 3), the integration can be further expressed in the following proposition.

PROPOSITION 2. *Let $f(s) = \mathbb{E}_t^{\mathbb{Q}}(\exp(sV_T)) = g(t, T-t, s)$, the VIX-based pricing formula for VIX call options is:*

$$C_t = e^{-r(T-t)}\left[f(1)P_1(t) - KP_2(t)\right]$$

where

$$\begin{aligned} P_1(t) &= \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty}\mathrm{Re}\left[\frac{f(i\phi+1)K^{-i\phi}}{i\phi f(1)}\right]d\phi, \\ P_2(t) &= \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty}\mathrm{Re}\left[\frac{f(i\phi)K^{-i\phi}}{i\phi}\right]d\phi, \end{aligned}$$

i and $\mathrm{Re}[x]$ are $\sqrt{-1}$ and the real part of x accordingly.

Proof. See appendix. □

As the $f(s) = g(t, T-t, s)$ is a function of lagged (log) volatility index $V_{t-1}, \ldots, V_{t+1-p}$ and innovations $\varepsilon_{t-1}, \ldots, \varepsilon_{t+1-q}$ (see Proposition 1), we refer the pricing formula in Proposition 2 as the VIX-based pricing formula.

3.2 Futures-based VIX option pricing formula

The formula provided in the previous section, to their best, still relies on VIX index levels. In this section, we start with the implied dynamics of VIX futures prices and then build a VIX futures based option pricing formula.

From Remark 1, the logarithm of model-implied VIX futures prices follows:

$$\log F_{t,t+k} = A(k) + \sum_{i=1}^{p} B_i(k)V_{t+1-i} + \sum_{i=1}^{q} C_i(k)\varepsilon_{t+1-i}$$

Define $\sigma_i = B_1(i) + C_1(i)$, the return of holding a particular VIX futures $F_{t,T}$ from time $t+j$ to time

$t+j+1$ can be written as[1]:

$$\log\left(\frac{F_{t+j+1,T}}{F_{t+j,T}}\right) = -\Psi(\sigma_{T-t-j-1}) + \sigma_{T-t-j-1}\varepsilon_{t+j+1}$$

By definition of $\Psi(x)$, the equation implies:

$$\mathbb{E}^{\mathbb{Q}}_{t+j}\left(\frac{F_{t+j+1,T}}{F_{t+j,T}}\right) = 1$$

which naturally satisfies the martingale condition for a tradable asset. The holding return then follows:

$$\begin{aligned}
\log\left(\frac{F_{T,T}}{F_{t,T}}\right) &= \sum_{j=0}^{k-1}\log\left(\frac{F_{t+j+1,T}}{F_{t+j,T}}\right) \\
&= \sum_{j=0}^{k-1}\left[-\Psi(\sigma_{k-j-1}) + \sigma_{k-j-1}\varepsilon_{t+j+1}\right] \\
&= \sum_{i=0}^{k-1}\left[-\Psi(\sigma_i) + \sigma_i\varepsilon_{t+k-i}\right] \\
&= \sum_{i=1}^{k}\left[-\Psi(\sigma_{i-1}) + \sigma_{i-1}\varepsilon_{t+k+1-i}\right]
\end{aligned}$$

where $k = T-t$. Note that $k-i \geq 0$ and ε_{t+j} is i.i.d. conditional on the information set on time t, the last equation is equivalent to:

$$\log(F_{T,T}) = \log(F_{t,T}) + \sum_{i=1}^{k}\left[-\Psi(\sigma_{i-1}) + \sigma_{i-1}\varepsilon_{t+i}\right]$$

Since σ_{i-1} are functions of parameters, the conditional moment generating function associated to $\log(F_{T,T}/F_{t,T})$ solely depends on the moment generating function of ε_{t+i}. The latter is previously defined as $\Psi(s)$. Therefore, the MGF of $\log(F_{T,T})$ is given by

$$\begin{aligned}
h(t,T-t,s) &= \mathbb{E}^{\mathbb{Q}}_t\left(\exp(s\log(F_{T,T}))\right) \\
&= \exp\left(s\log(F_{t,T}) + \sum_{i=1}^{T-t}\left[\Psi(s\sigma_{i-1}) - s\Psi(\sigma_{i-1})\right]\right) \qquad (1)
\end{aligned}$$

Using the fact that $\mathrm{VIX}_T = F_{T,T}$, the VIX call option price has the following representation:

$$\begin{aligned}
C_t &= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}_t\left[\max\left(\mathrm{VIX}_T - K, 0\right)\right] \\
&= e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}_t\left[\max\left(F_{T,T} - K, 0\right)\right]
\end{aligned}$$

Using similar technique for Proposition 2, the VIX option price can be linked to VIX futures through

[1] See appendix

the following formula.

PROPOSITION 3. *Let* $f^*(s) = \mathbb{E}_t^{\mathbb{Q}}(\exp(s\log(F_{T,T}))) = h(t, T-t, s)$*, the VIX futures based pricing formula for VIX call options is:*

$$C_t = e^{-r(T-t)}\left[f^*(1)\widetilde{P}_1(t) - K\widetilde{P}_2(t)\right]$$

where

$$\begin{aligned}\widetilde{P}_1(t) &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{f^*(i\phi+1)K^{-i\phi}}{i\phi f(1)}\right]d\phi, \\ \widetilde{P}_2(t) &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{f^*(i\phi)K^{-i\phi}}{i\phi}\right]d\phi,\end{aligned}$$

i and $\mathrm{Re}[x]$ *are* $\sqrt{-1}$ *and the real part of x accordingly.*

Proof. See appendix. □

Because $f^*(s) = h(t, T-t, s)$ is only a function of the prices of VIX futures (see Equation 1), we refer the pricing formula in Proposition 3 as the VIX futures based pricing formula. Almost for each VIX options, CBOE provides a VIX futures that matured at the same date. This provides us the ability to price VIX options using VIX futures through Proposition 3.

Note that although we express $f^*(s)$ as a function of VIX futures prices, theoretically it just quals to $f(s)$ defined by $g(t, T-t, s)$, as they are all MGF of $\log F_{T,T}$ (or V_T). However, this theoretical relationship may not always hold in reality. Therefore, deriving different forms of model-implied MGF helps us compare the information content from VIX index and VIX futures price, respectively.

4 Competing model

For discrete time framework, there are several models proposed for VIX option pricing. Cao et al. (2020) provides a way using Heston-Nandi GARCH model and inverse Gaussian distribution. Tong and Huang (2021) provides two models, the GARV and the Realized GARCH, using both S&P 500 index and realized volatility for option pricing. The former model is affine model while the latter model is non-affine model. In our empirical part, we compare our model with all these models mentioned above. As our model starts directly from the risk neutral measure, we introduce those models in their risk neutral forms too.

4.1 The Affine GARCH Models

The Heston-Nandi GARCH model

The Heston-Nandi GARCH model (Heston and Nandi, 2000) is one of the most widely used discrete time option pricing models. Its affine structure insures a semi-analytical solution for European call options price and such structure is found in most GARCH type option pricing models that yield close-from pricing formula.

The risk-neutral dynamics of the Heston-Nandi GARCH model follows:

$$\begin{aligned} R_{t+1} &= r - \tfrac{1}{2}h_{t+1} + \sqrt{h_{t+1}}z_{t+1} \\ h^*_{t+1} &= \omega^* + \beta h_t + \tau_2 \left(z_t - \tau_1\sqrt{h_t}\right)^2 \end{aligned}$$

where z_t follows standard normal distribution. The analytical VIX option pricing formula for the HNG model is provided in Cao et al. (2020). It is worth to mention that this model is a special case of the following Generalized Affine Realized Volatility model when the weight on the volatility component of return is one (i.e., $\xi = 1$).

The Generalized Affine Realized Volatility (GARV) model

Build on the Heston-Nandi GARCH model, Christoffersen et al. (2014) provided a generalized affine model with realized volatility (RV) augmented part to price index options. The risk neutral dynamics of GARV model follows:

$$\begin{aligned} R_{t+1} &= r - \tfrac{1}{2}h_{t+1} + \sqrt{h_{t+1}}z_{t+1} \\ h_{t+1} &= \xi h^R_{t+1} + (1-\xi)h^x_{t+1} \\ h^R_{t+1} &= \omega + \beta h^R_t + \tau_2(z_t - \tau_1\sqrt{h_t})^2 \\ h^x_{t+1} &= \kappa + \phi h^x_t + \delta_2(\varepsilon_t - \delta_1\sqrt{h_t})^2 \\ x_t &= h^x_t + \eta\left((\varepsilon_t - \delta_1\sqrt{h_t})^2 - 1 - \delta_1^2 h_t\right) \end{aligned}$$

where (z_t, ε_t) follows a standard bivariate normal distribution with the correlation of ρ. The key feature of this model is the conditional variance for returns, h_t, has two components. The first h^R_t, is driven by returns, and the second h^x_t, is driven by the realized measure x_t. The last equation is called measurement equation that links realized volatility to the conditional volatility. Note that the Heston-Nandi GARCH model (2000) is a special case of the GARV model when ξ is one.

The VIX option price for GARV model is provided in the Proposition 1 in Tong and Huang (2021) where the price is a integration of the moment generation function of forward expected variance.

$$C_t = \frac{e^{-r(T-t)}}{2\sqrt{\pi}} \int_0^\infty \Re\left[e^{ua} f^*_t(T,u) \times \frac{1-\mathrm{erf}(K\sqrt{u})}{(\sqrt{u})^3}\right] \mathrm{d}u_I$$

where $f^*_t(T,u)$ is the characteristic function of $y_T = bh^R_{T+1} + ch^x_{T+1}$:

$$f^*_t(T,u)\,\mathbb{E}^{\mathbb{Q}}_t(e^{uy_T}) = \exp\left(A(u,T) + B(u,T)h^R_{t+1} + D(u,T)h^x_{t+1}\right)$$

and u is a complex number denoted as $u = u_R + iu_I$, with $u_R > 0$ and $u_I \in \mathbb{R}$. $\Re[\cdot]$ stands for the real part of the complex number inside the square bracket. $\mathrm{erf}(x)$ is the error function defined as $\mathrm{erf}(x) :=$

$\frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-s^2} \mathrm{d}s$. Coefficients a, b, c are linked to model coefficients and $A(u,s)$, $B(u,s)$, $D(u,s)$ can be obtained through iterations using equations listed in the appendix of Tong and Huang (2021). This formula is also applicable for the Heston-Nandi GARCH model with proper constraints.

4.2 The Realized GARCH model

Unlike the previous two models, the Realized GARCH model proposed in Hansen and Huang (2016) is a non-affine model where the moment generation function of conditional forward variance is hard to obtain. In this case, Huang et al. (2017) applied the Edgeworth expansion to approximate the risk neutral distribution of cumulative returns for index option pricing. Expending option price with analytical approximation function enable us to calibrate parameters by minimizing pricing errors. The Monte-Carol simulation, while easier to apply, is not suitable for such calibration method due to the sampling error.

The risk neutral dynamics of the Realized GARCH model follows:

$$\begin{aligned} R_{t+1} &= r - \tfrac{1}{2} h_{t+1} + \sqrt{h_{t+1}} e_{t+1} \\ \log h_{t+1} &= \omega + \beta \log h_t + \tau_1 e_t + \tau_2 (e_t^2 - 1) + \gamma \sigma u_t \\ \log x_t &= \xi + \phi \log h_t + \delta_1 e_t + \delta_2 (e_t^2 - 1) + \sigma u_t \end{aligned}$$

The two innovations, e_t and u_t, are independent standard normal distribution. Define $\mu = \mathbb{E}_t^{\mathbb{Q}} (\log \mathrm{VIX}_T)$, $\sigma^2 = \mathrm{Var}_t^{\mathbb{Q}} (\log \mathrm{VIX}_T)$ the VIX call option price can be written as an integral of the variable $z_T = (\log \mathrm{VIX}_T - \mu)/\sigma$:

$$C_t = e^{-r(T-t)} \mathbb{E}_t^{\mathbb{Q}} (\max(\mathrm{VIX}_T - K, 0)) = e^{-r(T-t)} \int_{-\infty}^{k} [\exp(\mu - \sigma z) - K] \, \tilde{g}(z) \mathrm{d}z \qquad (2)$$

where $k = (-\log(K) + \mu)/\sigma$ and $\tilde{g}(z)$ is the true conditional density function of $-z_T$.

Jarrow and Rudd (1982) proposed a second order approximation of $\tilde{g}(z)$ using standard normal distribution and higher order moments of z:

$$\tilde{g}(z) \approx \left[1 - \frac{\kappa_3}{6} H_3(z) + \frac{(\kappa_4 - 3)}{24} H_4(z) + \frac{\kappa_3^2}{72} H_6(z) \right] \phi(z), \qquad (3)$$

where $\phi(z)$ is the density function of the standard normal distribution, κ_i is the i-th moment of z_T, and $H_n(z)$ is the n-th order Hermite polynomial. Tong and Huang (2021) provide the expansion based pricing formula in their Proposition 2. The European VIX call price equals:

$$C_{approx} = A_0 - \frac{\kappa_3}{6} A_3 + \frac{(\kappa_4 - 3)}{24} A_4 + \frac{\kappa_3^2}{72} A_6 \qquad (4)$$

where A_0, A_3, A_4, A_6 are defined with moments of normal distribution and κ_i are defined as the standardized origin moments of $\log \mathrm{VIX}_t$. Exact formula for $\mathbb{E}_t (\log \mathrm{VIX}_T)^i$ is hard to compute and Tong and Huang (2021) provided a approximation for those moments through Taylor expansion.

5 Empirical Results

5.1 Data

Our dataset contains S&P 500 index from Yahoo Finance, realized volatility is obtained from the Realized Library of Oxford-Man institute[2], VIX index and VIX futures are downloaded from CBOE website, and VIX option price data are obtained from the CBOE DataShop. The full sample spans the periods from February 1, 2006 to October 27, 2020, with 3,711 trading days. For each options, CBOE provides a corresponding VIX futures with the same maturity.

For liquidity consideration, following Song and Xiu (2016) and Luo et al. (2019), the option data is trimmed using the following filter: 1) Options with time to maturity less than 7 or more than 126 days are dropped. 2) Mid-quote less than 0.1 or the relative spread[3] is greater than 0.3. In addition, only call options are used as major trading volume of VIX options concentrates on calls. A rolling window out-of-sample comparison with a window length of 10 years are also included.

The summary statistics are listed in Table 1. As shown in Table 1, compared with realized volatility, the CBOE VIX has higher mean which is in line with literature on variance risk premium. The standard deviation and kurtosis of VIX are also lower than realized variance indicating that VIX is less volatile than corresponding realized volatility. For VIX options, we documented implied volatility smirk and downward slop term structure.

[Insert Table 1 here]

5.2 Calibration

As we only focused on risk neutral dynamics, the parameters are calibrated by matching the model price with corresponding market price by minimizing the sum square of pricing errors. Given a set of parameters Θ, we can calculate the corresponding option price using VIX based method ($C_{t,T}^{Mod}(\Theta|VIX_t)$) and future based method ($C_{t,T}^{Mod}(\Theta|F_{t,T})$). The pricing error is then defined as

$$e_i(\Theta) = C_i^{Mkt} - C_i^{Mod}(\Theta)$$

where "Mkt" and "Mod" denote market price and model price, respectively. Following Yin et al. (2021) and others, we assume that the pricing errors are independent and normally distributed with mean zero and variance σ_e^2 . The corresponding log-likelihood function is then specified as:

$$\ell_o = -\frac{1}{2}\sum_{t=1}^{T}\sum_{i=1}^{N_t}\left\{\log\left(2\pi\sigma_e^2\right) + \frac{e_i^2(\Theta)}{\sigma_e^2}\right\}$$

where T is the number of trading days, and N_t is the number of VIX option prices on the day t. σ_e^2 can be estimated with the sample variance of pricing errors according to the first-order conditions

[2]Realized volatility in this paper is calculated using the realized kernel method to mitigate the impact of market micro-structure noise.

[3]Defined as (offer - bid)/mid-quote where mid-quote = (offer+bid)/2.

in the likelihood estimation. The calibration method is also applied to our competing model with information set $C_{t,T}^{Mod}(\Theta|R_t)$ and $C_{t,T}^{Mod}(\Theta|R_t,RV_t)$ for Heston-Nandi GARCH and GARV/Realized GARCH respectively.

We summarize the calibration results for ARMA type models in Table 2. For ARMA type models, we use three different settings including ARMA(1,1), HAR up to monthly lags (HAR-M) and HAR up to yearly lags (HAR-Y). The last setup is motivated by Huang et al. (2019) which suggested the importance of quarterly and yearly lags. For each setting, we discuss the difference when jumps are included and the shift of pricing strategies.

[Insert Table 2 here]

First, for all settings, the persistence parameters $\pi^{\mathbb{Q}}$ are highly close to one indicating high persistence for logVIX series. For lag parameters, in line with Yin et al. (2021), daily lags receive the largest weights while weekly parameters are the either smaller or close-to-zero compared with other lag parameters. For random shocks, judging by the difference in log-likelihoods, the jump components are statistically significant . Asymmetric jump features are also documented as the upward jumps dominated ($p_{\text{up}} > 0.8$) most jump cases. The average jump size for upward jumps is also much larger than downward jumps. For futures based pricing strategy, the calibrated jumps intensity is much higher, the proportion of upward jumps are higher, the difference in jump size between upward and downward is lower. The variance of continuous component δz_t is naturally higher when jumps are eliminated.

[Insert Table 3 here]

Parameters for competing GARCH models are listed in Table 3. Results are all in line with previous literature with strong leverage effect and high persistence.

5.3 Pricing performance

Following Tong and Huang (2021) and Jing et al. (2020), the pricing performance is evaluated through the root mean square error:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(C_i^{Mkt} - C_i^{Mod})^2}$$

where Mkt and Mod indicates market price and model price receptively. We also provide sub-sample RMSE with respect to different moneyness level (defined by $\log(F/K)$) and time to maturity. In line with previous section, ARMA(1,1), HAR with two different lag specifications, and three different SPX-based models are discussed.

5.3.1 In-sample

In sample results are listed in Table 4. For full sample RMSE, the best set of models is futures-based models and than the VIX-based models. The SPX-based models ranks the last.

[Insert Table 4 here]

Interestingly, adding jumps can significantly improve model fit for future-based models while the improvements are marginal for VIX-based models. On the other side, the long memory features make little difference for futures-based models while the reduction is important for VIX-based models. Both results for VIX-based models are in line with Yin et al. (2021) while the futures-based results shows strikingly difference in which settings are favorable. We provide a figure reports the reduction of RMSE due to difference settings in Figure 1.

[Insert Figure 1 here]

In Figure 1, we report two layers of RMSE reduction. The first layer focus on information set used and the corresponding base case (Ret Based) is defined as the RMSE of the worse performed Heston-Nandi GARCH model. The second layer focus on model specifications such as log-linear alternations or jump augmentations. For the second layer, the base cases are the corresponding first layer RMSE. For example, the Ret RV based model in this paper is the GARV model while its log-linear specification is defined as Realized GARCH model. As multiple specifications are discussed in this paper, we average the RMSE of ARMA(1,1), HAR(M) and HAR(Y) when possible.

In line with Tong and Huang (2021), adding RV into the GARCH framework can reduce RMSE significantly. In our case, the linear GARV model reduces RMSE by 25.72% and a further 14.14% reduction can be achieved when we switch to log-linear RG model. Direct strategy reduces RMSE much greater than indirect strategies with realized variance especially when futures are used. Even the vanilla ARMA version with futures information can reduce RMSE by 67.9% and a further 7.15% reduction is reported with jumps. With proper information, our simple framework can reduce over 75% of in-sample RMSE over the benchmark Heston-Nandi GARCH model and nearly 60% of in-sample RMSE over log-linear RG model. The decomposed results confirms those finding in full RMSE. A interesting finding here is that the RMSE patterns over moneyness and maturity are rather flat for future-based models while upward sloped patterns are reported for other models.

To sum up, our in-sample findings suggest that futures that share identical time to maturity of options contains much richer information than VIX index itself, let alone the SP500 index and its realized volatility.

5.3.2 Out-of-sample

A common concern for the in-sample winner is that whether we push the model too hard and overfit the data. Judging by the model structure, our framework is stronger in resisting overfit as our models are simpler than SPX-based models. We improvement pricing utilizing option specific information embedded in the corresponding futures rather than flooding the model with tons of parameters. Nevertheless, we provide out-of-sample pricing evaluation with two different setting in this subsection. A "estimate-and-forget" method estimates parameters ones with the first ten years data (2006 to 2015) and then use them to price options for the rest years (2016-2020). A "rolling-window" method updates parameters every month since 2016 using a ten year estimation window. Results are summarized in Table 5 and Table 6 respectively. Similar to in-sample results, we summarize RMSEs for different models, different information sets and decompose options into moneyness as well as time-to-maturity groups.

[Insert Table 5 here]

[Insert Table 6 here]

[Insert Figure 2 here]

We also calculate two layers of RMSE reductions and plot them in Figure 2. As two out-of-sample settings are used, we report "estimate-and-forget" results in parenthesis and "rolling-window" results in brackets. Although the exact numbers differs, the out-of-sample RMSE reductions are not only similar to each other but also close to in-sample RMSE reductions. These suggest low possibility of in-sample overfit. We still conclude that switching from SPX information to VIX/VIX futures can significantly improve model performance. Utilizing information from option specific futures, one can reduce RMSE further 32% in addition to the log-linear RG model (the best performed SPX-based model in this paper).

6 Conclusion

In this paper, we use a simple ARMA framework with jumps to model log-VIX dynamics and provide explicit links between VIX index, VIX futures and VIX options. By doing this, two explicit VIX option pricing formulas are derived to explored the pricing implication due to the fact that each VIX option has a corresponding VIX futures with same time-to-maturity. We provided extensive empirical evidence based on CBOE VIX options from 2006 to 2020 that support the new framework's significant performance gains over existing frameworks based on SPX daily returns, realized variance, or VIX index itself. Among these models, the futures-based model provides the best pricing performance, with a reduction in RMSE up to 50% compared with the VIX-based model and a reduction up to nearly 60% compared with the SPX-based model including realized volatility.

References

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003, March). Modeling and Forecasting Realized Volatility. *Econometrica 71*(2), 579–625.

Cao, H., A. Badescu, Z. Cui, and S. K. Jayaraman (2020). Valuation of VIX and target volatility options with affine GARCH models. *Journal of Futures Markets 40*(12), 1880–1917.

Cao, J., X. Ruan, S. Su, and W. Zhang (2020). Pricing VIX derivatives with infinite-activity jumps. *Journal of Futures Markets 40*(3), 329–354.

Christoffersen, P., B. Feunou, K. Jacobs, and N. Meddahi (2014). The economic value of realized volatility: Using high-frequency returns for option valuation. *Journal of Financial and Quantitative Analysis 49*, 663–697.

Detemple, J. and C. Osakwe (2000, 04). The Valuation of Volatility Options. *Review of Finance 4*(1), 21–50.

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

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

Hansen, P. R. and Z. Huang (2016). Exponential GARCH modeling with realized measures of volatility. *Journal of Business & Economic Statistics 34*, 269–287.

Heston, S. L. and S. Nandi (2000). A closed-form GARCH option valuation model. *Review of Financial Studies 13*, 585–625.

Huang, Z., C. Tong, and T. Wang (2019). VIX term structure and VIX futures pricing with realized volatility. *Journal of Futures Markets 39*(1), 72–93.

Huang, Z., T. Wang, and P. R. Hansen (2017). Option pricing with the realized GARCH model: An analytical approximation approach. *Journal of Futures Markets 37*, 328–358.

Jarrow, R. and A. Rudd (1982). Approximate option valuation for arbitrary stochastic processes. *Journal of Financial Economics 10*(3), 347–369.

Jing, B., S. Li, and Y. Ma (2020). Pricing VIX options with volatility clustering. *Journal of Futures Markets 40*(6), 928–944.

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

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.

Lin, Y.-N. (2007). Pricing VIX futures: Evidence from integrated physical and risk-neutral probability measures. *Journal of Futures Markets 27*(12), 1175–1217.

Lin, Y.-N. (2013). Vix option pricing and cboe vix term structure: A new methodology for volatility derivatives valuation. *Journal of Banking & Finance 37*(11), 4432–4446.

Luo, X., J. E. Zhang, and W. Zhang (2019). Instantaneous squared VIX and VIX derivatives. *Journal of Futures Markets 39*(10), 1193–1213.

Mencía, J. and E. Sentana (2013). Valuation of VIX derivatives. *Journal of Financial Economics 108*(2), 367–391.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica 59*, 347–370.

Park, Y.-H. (2016). The effects of asymmetric volatility and jumps on the pricing of VIX derivatives. *Journal of Econometrics 192*(1), 313–328.

Psychoyios, D., R. Dotsis, and R. N. Markellos (2010). A jump diffusion model for VIX volatility options and futures. *Review of Quantitative Finance & Accounting 35*(3), 245–269.

Sepp, A. (2008). VIX option pricing in a jump-diffusion model. *Risk Magazine*, 84–89.

Song, Z. and D. Xiu (2016). A tale of two option markets: Pricing kernels and volatility risk. *Journal of Econometrics 190*, 176–196.

Tong, C. and Z. Huang (2021). Pricing VIX options with realized volatility. *Journal of Futures Markets 41*(8), 1180–1200.

Wang, Q. and Z. Wang (2021). VIX futures and its closed-form pricing through an affine GARCH model with realized variance. *Journal of Futures Markets 41*(1), 135–156.

Wang, T., Y. Shen, Y. Jiang, and Z. Huang (2017). Pricing the CBOE VIX futures with the Heston–Nandi GARCH model. *Journal of Futures Markets 37*(7), 641–659.

Yang, X. (2018). Good jump, bad jump, and option valuation. *Journal of Futures Markets 38*(9), 1097–1125.

Yang, X. and P. Wang (2018). VIX futures pricing with conditional skewness. *Journal of Futures Markets 38*(9), 1126–1151.

Yang, X., P. Wang, and J. Chen (2019). VIX futures pricing with affine jump-GARCH dynamics and variance-dependent pricing kernels. *The Journal of Derivatives 27*(1), 110–127.

Yin, F., Y. Bian, and T. Wang (2021). A short cut: Directly pricing vix futures with discrete-time long memory model and asymmetric jumps. *Journal of Futures Markets 41*(4), 458–477.

Zhang, J. E., J. Shu, and M. Brenner (2010). The new market for volatility trading. *Journal of Futures Markets 30*(9), 809–833.

Zhang, J. E. and Y. Zhu (2006). VIX futures. *Journal of Futures Markets 26*(6), 521–531.

Zhu, S.-P. and G.-H. Lian (2012). An analytical formula for VIX futures and its applications. *Journal of Futures Markets 32*(2), 166–190.

A Appendix of Proofs

Our model for $V_t = \log(\mathrm{VIX}_t)$ follows an ARMA(p,q) process with jumps:

$$V_t = \mu + \sum_{i=1}^{p} \beta_i V_{t-i} + \sum_{j=1}^{q} \alpha_i \varepsilon_{t-j} + \varepsilon_t$$

and define

$$\begin{aligned} \varepsilon_t &\equiv \delta z_t + J_t \\ \Psi(s) &\equiv \log \mathbb{E}_t(\exp(s\varepsilon_{t+1})) \end{aligned}$$

In our following empirical analysis, we use the asymmetric jump structure

$$z_t \sim N(0,1), \quad J_t = \sum_{j=0}^{N_t} e_{j,t}, \quad e_{j,t} \sim \text{Kou}(p,q,\eta_u,\eta_d), \quad N_t \sim Poisson(\lambda)$$

B Proof of PROPOSITION 1

Let $k = T - t$, and assume that

$$\mathbb{E}_t^{\mathbb{Q}}\left(\exp\left(sV_{t+k}\right)\right) = \exp\left(A(k,s) + \sum_{i=1}^{p} B_i(k,s)V_{t+1-i} + \sum_{j=1}^{q} C_i(k,s)\varepsilon_{t+1-j}\right)$$

For $k = 0$, we have

$$A(0,s) = 0, \quad B_i(0,s) = \begin{cases} s & i = 1 \\ 0 & 1 < i < p \end{cases}, \quad C(0,s) = 0$$

For $k = k+1$, we have

$$\begin{aligned} \mathbb{E}_t^{\mathbb{Q}}\left(\exp\left(sV_{t+k+1}\right)\right) &= \mathbb{E}_t^{\mathbb{Q}}\left(\exp\left(A(k,s) + \sum_{i=1}^{p} B_i(k,s)V_{t+2-i} + \sum_{i=1}^{q} C_i(k,s)\varepsilon_{t+2-i}\right)\right) \\ &= \exp\left(A(k,s) + \sum_{i=2}^{p} B_i(k,s)V_{t+2-i} + \sum_{i=2}^{q} C_i(k,s)\varepsilon_{t+2-i}\right) \\ &\quad \times \mathbb{E}_t^{\mathbb{Q}}\left(\exp\left(B_1(k,s)V_{t+1} + C_1(k,s)\varepsilon_{t+1}\right)\right) \\ &= \exp\left(A(k,s) + \sum_{i=1}^{p-1} B_{i+1}(k,s)V_{t+1-i} + \sum_{i=1}^{q-1} C_{i+1}(k,s)\varepsilon_{t+1-i}\right) \\ &\quad \times \exp\left(B_1(k,s)\mu + \sum_{i=1}^{p} B_1(k,s)\beta_i V_{t+1-i} + \sum_{i=1}^{q} B_1(k,s)\alpha_i\varepsilon_{t+1-i} + \Psi(B_1(k,s) + C_1(k,s))\right) \\ &= \exp\left(A(k+1,s) + \sum_{i=1}^{p} B_i(k+1,s)V_{t+1-i} + \sum_{i=1}^{q} C_i(k+1,s)\varepsilon_{t+1-i}\right) \end{aligned}$$

with

$$\begin{aligned} A(k+1,s) &= A(k,s)+B_1(k,s)\mu+\Psi(B_1(k,s)+C_1(k,s)) \\ B_i(k+1,s) &= \begin{cases} B_{i+1}(k,s)+B_1(k,s)\beta_i & 1\le i<p \\ B_1(k,s)\beta_i & i=p \end{cases} \\ C_i(k+1,s) &= \begin{cases} C_{i+1}(k,s)+B_1(k,s)\alpha_i & 1\le i<q \\ B_1(k,s)\alpha_i & i=q \end{cases} \end{aligned}$$

Therefore, we have the model implied VIX futures, given by

$$F_{t,T} = \mathbb{E}_t^{\mathbb{Q}}\left(\exp\left(V_{t+k}\right)\right) = \exp\left(A(k,1)+\sum_{i=1}^{p} B_i(k,1)V_{t+1-i}+\sum_{i=1}^{q} C_i(k,1)\varepsilon_{t+1-i}\right)$$

C Futures-based VIX option pricing formula

C.1 Model-implied VIX futures

In the previous, we have derived the formula of $\mathbb{E}_t^{\mathbb{Q}}\left(\exp\left(sV_{t+k}\right)\right)$, so we have

$$F_{t,T} = \mathbb{E}_t^{\mathbb{Q}}\left(\exp\left(V_{t+k}\right)\right) = \exp\left(A(k)+\sum_{i=1}^{p} B_i(k)V_{t+1-i}+\sum_{i=1}^{q} C_i(k)\varepsilon_{t+1-i}\right)$$

$$A(0)=0, \quad C(0)=0, \quad B_i(0)=\begin{cases} 1 & i=1 \\ 0 & 1<i<p \end{cases}$$

$$\begin{aligned} A(k+1) &= A(k)+B_1(k)\mu+\Psi(B_1(k)+C_1(k)) \\ B_i(k+1) &= \begin{cases} B_{i+1}(k)+B_1(k)\beta_i & 1\le i<p \\ B_1(k)\beta_i & i=p \end{cases} \\ C_i(k+1) &= \begin{cases} C_{i+1}(k)+B_1(k)\alpha_i & 1\le i<q \\ B_1(k)\alpha_i & i=q \end{cases} \end{aligned}$$

C.2 The Dynamic of VIX Futures Prices

So, the (log) model implied VIX futures price is given by

$$\log F_{t,t+k} = A(k)+\sum_{i=1}^{p} B_i(k)V_{t+1-i}+\sum_{i=1}^{q} C_i(k)\varepsilon_{t+1-i}$$

Let $k = T - t$, $\sigma_i = B_1(i) + C_1(i)$, we have

$$\begin{aligned}
\log\left(\frac{F_{t+j+1,T}}{F_{t+j,T}}\right) &= A(k-j-1) - A(k-j) + \sum_{i=1}^{p} B_i(k-j-1)V_{t+j+2-i} - \sum_{i=1}^{p} B_i(k-j)V_{t+j+1-i} \\
&\quad + \sum_{i=1}^{q} C_i(k-j-1)\varepsilon_{t+j+2-i} - \sum_{i=1}^{q} C_i(k-j)\varepsilon_{t+j+1-i} \\
&= -B_1(k-j-1)\mu - \Psi(\psi(k-j-1)) + \sum_{i=1}^{p-1} B_{i+1}(k-j-1)V_{t+j+1-i} \\
&\quad + B_1(k-j-1)(\mu + \sum_{i=1}^{p} \beta_i V_{t+j+1-i} + \sum_{i=1}^{q} \alpha_i \varepsilon_{t+j+1-i} + \varepsilon_{t+j+1}) - \sum_{i=1}^{p} B_i(k-j)V_{t+j+1-i} \\
&\quad + \sum_{i=1}^{q-1} C_{i+1}(k-j-1)\varepsilon_{t+j+1-i} + C_1(k-j-1)\varepsilon_{t+j+1} - \sum_{i=1}^{q} C_i(k-j)\varepsilon_{t+j+1-i} \\
&= -\Psi(B_1(k-j-1) + C_1(k-j-1)) + (B_1(k-j-1) + C_1(k-j-1))\varepsilon_{t+j+1} \\
&= -\Psi(\sigma_{k-j-1}) + \sigma_{k-j-1}\varepsilon_{t+j+1}
\end{aligned}$$

Note that it satisfies following martingale condition for a tradable future asset:

$$\mathbb{E}^{\mathbb{Q}}_{t+j}\left(\frac{F_{t+j+1,T}}{F_{t+j,T}}\right) = 1$$

Then we have

$$\begin{aligned}
R^{\text{Fut}}_{t,T} \equiv \log\left(\frac{F_{T,T}}{F_{t,T}}\right) &= \sum_{j=0}^{k-1} \log\left(\frac{F_{t+j+1,T}}{F_{t+j,T}}\right) \\
&= \sum_{j=0}^{k-1} \left[-\Psi(\sigma_{k-j-1}) + \sigma_{k-j-1}\varepsilon_{t+j+1}\right] \\
&= \sum_{i=0}^{k-1} \left[-\Psi(\sigma_i) + \sigma_i \varepsilon_{t+k-i}\right] \\
&= \sum_{i=1}^{k} \left[-\Psi(\sigma_{i-1}) + \sigma_{i-1}\varepsilon_{t+k+1-i}\right]
\end{aligned}$$

Table 1: Summary Statistics

A: SPX returns, Realized Volatility, and CBOE VIX (2006-2020)

	Mean(%)	Std(%)	Skewness	Kurtosis	Obs.
Returns (annualized)	7.703	17.846	-0.554	16.495	3,711
Realized Volatility (annualized)	12.495	10.502	3.493	21.281	3,711
CBOE VIX	19.479	9.649	2.442	11.050	3,711

B: VIX Option Price Data (Wednesday: 2006-2020)

	Average Price ($)	Implied Volatility	Obs.
All VIX call options	1.607	1.001	50,945
Partitioned by Moneyness, $m := \log(F/K)$			
m<-0.4	0.472	1.205	19,512
-0.4 $\leq m$<-0.2	1.059	1.022	10,372
-0.2$\leq m$<-0.1	1.594	0.925	4,851
-0.1$\leq m$<0	2.107	0.852	4,544
0$\leq m$<0.1	2.759	0.790	4,030
0.1$\leq m$	4.352	0.696	7636
Partitioned by Days to Maturity			
DTM<20	1.017	1.277	7,414
20≤DTM<40	1.353	1.120	10,323
40≤DTM<60	1.624	0.994	6,059
60≤DTM<80	1.741	0.932	7,967
80≤DTM<100	1.905	0.861	7,316
100≤DTM	2.055	0.791	7,978

Note: Summary statistics for close-to-close S&P 500 index returns, realized kernels (in square root), CBOE VIX and VIX option prices from February 1, 2006 to October 27, 2020. The reported statistics for S&P 500, realized kernels, and VIX index include the sample mean (Mean), standard deviation (Std), skewness (Skew), kurtosis (Kurt), number of observations (Obs). We report the average price, average implied volatility, and the number of option prices for different partitions of our (Wednesday) VIX option prices. "Moneyness" is defined by the $m = \log F/K$, where F the is VIX futures price and K is the strike price. DTM denotes the number of days to maturity. Data sources: S&P 500 returns from Yahoo Finance; VIX and VIX futures from CBOE's website; Realized kernels from Realized Library of Oxford-Man institute; Option prices from CBOE Data Shop.

Table 2: Estimation Results for VIX-based and Futures-based Models

	VIX-based Models						Futures-based Models					
	With Jump			No Jump			With Jump			No Jump		
Model	ARMA(1,1)	HAR-M	HAR-Y	ARMA(1,1)	HAR-M	HAR-Y	ARMA(1,1)	HAR-M	HAR-Y	ARMA(1,1)	HAR-M	HAR-Y
μ	0.0150	0.0188	0.0084	0.0143	0.0184	0.0089						
	(0.0003)	*(0.0003)*	*(0.0003)*	*(0.0002)*	*(0.0003)*	*(0.0003)*						
β_d	0.9939	0.9445	0.9674	0.9939	0.9541	0.9742	0.9912	0.9332	0.9468	0.9927	0.9332	0.9477
	(0.0001)	*(0.0002)*	*(0.0001)*	*(0.0001)*	*(0.0002)*	*(0.0001)*	*(0.0003)*	*(0.0329)*	*(0.0364)*	*(0.0005)*	*(0.0305)*	*(0.0796)*
β_w		0.0035	0.0042		0.0002	0.0008		0.0026	0.0026		0.0023	0.0010
		(0.0011)	*(0.0009)*		*(0.0009)*	*(0.0009)*		*(0.0312)*	*(0.0353)*		*(0.0292)*	*(0.0794)*
β_m		0.0441	0.0091		0.0382	0.0078		0.0509	0.0314		0.0510	0.0324
		(0.0010)	*(0.0010)*		*(0.0009)*	*(0.0010)*		*(0.0020)*	*(0.0035)*		*(0.0021)*	*(0.0046)*
β_q			0.0105			0.0089			0.0122			0.0109
			(0.0003)			*(0.0003)*			*(0.0026)*			*(0.0027)*
β_y			0.0053			0.0048			0.0025			0.0016
			(0.0001)			*(0.0001)*			*(0.0870)*			*(0.0675)*
α	-0.3468			-0.3009			-0.6229			-0.6912		
	(0.0446)			*(0.0366)*			*(0.0100)*			*(0.0122)*		
δ	0.1141	0.0898	0.0655	0.1234	0.0988	0.0963	0.0659	0.0488	0.0467	0.1824	0.0999	0.0948
	(0.0162)	*(0.0013)*	*(0.0013)*	*(0.0057)*	*(0.0005)*	*(0.0005)*	*(0.0027)*	*(0.0029)*	*(0.0031)*	*(0.0040)*	*(0.0053)*	*(0.0127)*
p_{up}	0.9191	0.8607	0.8081				0.8981	0.8488	0.8481			
	(0.0777)	*(0.0392)*	*(0.0461)*				*(0.0671)*	*(0.0033)*	*(0.0639)*			
$1/\eta_{\mathrm{up}}$	0.9450	0.4395	0.4088				0.3824	0.3037	0.2927			
	(0.1377)	*(0.0193)*	*(0.0128)*				*(0.0082)*	*(0.0186)*	*(0.0188)*			
$1/\eta_{\mathrm{down}}$	0.0023	0.0023	0.0023				0.0023	0.0023	0.0023			
	(0.0033)	*(0.0056)*	*(0.0061)*				*(0.0005)*	*(0.0010)*	*(0.0049)*			
λ	0.0004	0.0068	0.0083				0.0589	0.0340	0.0328			
	(0.0007)	*(0.0014)*	*(0.0016)*				*(0.0029)*	*(0.0019)*	*(0.0028)*			
$\pi^{\mathbb{Q}}$	0.9939	0.9921	0.9965	0.9939	0.9925	0.9965	0.9912	0.9867	0.9954	0.9927	0.9865	0.9937
ℓ	-58019	-54027	-45173	-58414	-54874	-47446	-16470	-16330	-16297	-29315	-29293	-29283

Note: Estimation results for VIX-based and Futures-based models in the full sample period (February 1, 1990 to October 27, 2020) . Parameter estimates are reported with robust standard errors (in parentheses), $\pi^{\mathbb{Q}}$ refer to the volatility persistence under risk-neutral measures. The value of the log-likelihood function is reported at the bottom of the table.

Table 3: Estimation Results for SPX-based Models

	HNG	GARV	RG
β	0.2765	0.9880	0.9974
	(0.0410)	*(0.0090)*	*(0.0054)*
τ_1	765.87	116.10	-0.0197
	(20.10)	*(19.77)*	*(0.002)*
τ_2	1.23E-06	8.73E-07	0.0055
	(7.51E-08)	*(6.85E-08)*	*(0.001)*
γ		0.0857	0.1847
		(0.0131)	*(0.0102)*
ξ		0.0189	-0.5643
		(0.0043)	*(0.0675)*
ϕ		3.74E-08	1.0340
		(1.67E-03)	*(0.023)*
δ_1		888.60	-0.0687
		(25.70)	*(0.001)*
δ_2		1.27E-06	0.1748
		(3.90E-08)	*(0.036)*
σ/ρ		0.9986	0.8950
		(0.0330)	*(0.0216)*
$\log \mathbb{E}^{\mathbb{Q}}(h_t)$	-8.1587	-8.3020	-8.2378
	(0.1100)	*(0.0810)*	*(0.0698)*
$\pi^{\mathbb{Q}}$	0.9957	0.9998	0.9974
ℓ	-86002	-74238	-65030

Note: Estimation results for SPX-based models in the full sample period (February 1, 1990 to October 27, 2020) . Parameter estimates are reported with robust standard errors (in parentheses), $\pi^{\mathbb{Q}}$ refer to the volatility persistence under risk-neutral measures. The value of the log-likelihood function is reported at the bottom of the table.

Table 4: In-sample VIX Options Pricing Performance

	VIX-based Models						Futures-based Models						SPX-based Models		
	With Jump			No Jump			With Jump			No Jump					
Model	ARMA11	HAR(M)	HAR(Y)	ARMA11	HAR(M)	HAR(Y)	ARMA11	HAR(M)	HAR(Y)	ARMA11	HAR(M)	HAR(Y)	HNG	GARV	RG
Full RMSE	0.7557	0.6988	0.5873	0.7616	0.7105	0.6141	0.3343	0.3334	0.3332	0.4302	0.4300	0.4299	1.3403	0.9956	0.8065
Part 1: Evaluation By Moneyness, m=log(F/K)															
$[-\infty,-0.4]$	0.3182	0.2832	0.2785	0.3424	0.3137	0.3300	0.2388	0.2386	0.2386	0.3744	0.3741	0.3742	0.5120	0.5272	0.3422
$(-0.4,-0.2]$	0.5736	0.5091	0.4862	0.5971	0.5389	0.5189	0.3753	0.3749	0.3748	0.4440	0.4437	0.4434	0.9475	0.8395	0.6292
$(-0.2,-0.1]$	0.7482	0.6755	0.6201	0.7645	0.7129	0.6594	0.4175	0.4168	0.4167	0.4248	0.4244	0.4240	1.3233	1.0305	0.8039
$(-0.1,0]$	0.8628	0.7931	0.6984	0.8663	0.8219	0.7403	0.4376	0.4363	0.4357	0.4553	0.4554	0.4551	1.6010	1.2240	0.9212
$(0,0.1]$	0.9854	0.9259	0.7736	0.9720	0.9329	0.7989	0.4209	0.4184	0.4179	0.5100	0.5097	0.5098	1.9345	1.3826	1.0514
$(0.1,\infty]$	1.3330	1.2525	0.9607	1.3240	1.2342	0.9616	0.3008	0.2985	0.2981	0.4823	0.4824	0.4825	2.3417	1.5394	1.4153
Part 2: Evaluation By Days to Maturity (DTM)															
$[0,20]$	0.4338	0.3776	0.3659	0.4424	0.4034	0.3997	0.2942	0.2902	0.2901	0.3571	0.3564	0.3562	1.1026	0.8186	0.8717
$(20,40]$	0.6042	0.5637	0.4885	0.6135	0.5787	0.5190	0.3260	0.3257	0.3253	0.4133	0.4128	0.4127	1.2488	0.9368	0.8119
$(40,60]$	0.8246	0.7596	0.6730	0.8290	0.7673	0.6932	0.3324	0.3322	0.3320	0.4391	0.4391	0.4390	1.4374	1.0896	0.8563
$(60,80]$	0.8246	0.7721	0.6407	0.8299	0.7823	0.6647	0.3347	0.3345	0.3342	0.4424	0.4424	0.4423	1.4191	1.0475	0.7748
$(80,100]$	0.8767	0.8122	0.6564	0.8810	0.8224	0.6840	0.3385	0.3382	0.3380	0.4471	0.4471	0.4471	1.4129	1.0416	0.7646
$(100,\infty]$	0.9040	0.8334	0.6714	0.9094	0.8432	0.7003	0.3743	0.3733	0.3733	0.4758	0.4758	0.4758	1.4262	1.0461	0.7657

Note: This table reports the in-sample VIX option pricing performance for each model from February 1, 2006 to October 27, 2020. We evaluate the model's VIX option pricing ability through the root of mean square errors of option prices (RMSE). We summarize the results by option moneyness, and days to maturity. "Moneyness" is defined by the $m = \log F/K$, where F is the VIX futures price and K is the strike price. DTM denotes the number of days to maturity.

Table 5: Out-of-sample VIX Options Pricing Performance (Estimate Once)

	VIX-based Models						Futures-based Models						SPX-based Models		
	With Jump			No Jump			With Jump			No Jump					
Model	ARMA11	HAR(M)	HAR(Y)	ARMA11	HAR(M)	HAR(Y)	ARMA11	HAR(M)	HAR(Y)	ARMA11	HAR(M)	HAR(Y)	HNG	GARV	RG
Full RMSE	0.8151	0.7451	0.6692	0.8203	0.7591	0.6994	0.4193	0.4184	0.4331	0.5068	0.5065	0.5170	1.4892	1.0768	0.9057
Part 1: Evaluation By Moneyness m=log(F/K)															
$[-\infty, -0.4]$	0.3273	0.2830	0.3170	0.3636	0.3476	0.4153	0.3114	0.3144	0.3252	0.4799	0.4798	0.4890	0.5761	0.5712	0.3734
$(-0.4, -0.2]$	0.5902	0.5312	0.5413	0.6098	0.5604	0.5917	0.4820	0.4769	0.4947	0.6300	0.6294	0.6471	1.1084	0.9041	0.7844
$(-0.2, -0.1]$	0.7518	0.6924	0.6564	0.7616	0.7092	0.6644	0.5341	0.5301	0.5489	0.5680	0.5675	0.5844	1.5657	1.1092	1.0341
$(-0.1, 0]$	0.8894	0.8136	0.7228	0.8932	0.8253	0.7181	0.5657	0.5631	0.5826	0.5046	0.5049	0.5188	1.8794	1.3058	1.1978
$(0, 0.1]$	1.1174	1.0306	0.9118	1.1088	1.0244	0.8970	0.5508	0.5504	0.5710	0.4562	0.4557	0.4644	2.2344	1.5400	1.3669
$(0.1, \infty]$	1.5320	1.4093	1.1877	1.5169	1.3974	1.1920	0.3564	0.3567	0.3669	0.3731	0.3731	0.3664	2.6031	1.7593	1.4403
Part 2: Evaluation By Days to Maturity (DTM)															
$[0, 20]$	0.4336	0.3388	0.3740	0.4330	0.3539	0.3806	0.3450	0.3348	0.3360	0.4099	0.4100	0.4109	1.1545	0.8319	0.9874
$(20, 40]$	0.6229	0.5706	0.5264	0.6247	0.5799	0.5411	0.4021	0.4024	0.3994	0.4892	0.4874	0.4856	1.3804	1.0061	0.8602
$(40, 60]$	0.9541	0.8855	0.8350	0.9572	0.8939	0.8527	0.4136	0.4205	0.4221	0.5266	0.5278	0.5284	1.6910	1.2597	1.0094
$(60, 80]$	0.8988	0.8462	0.7526	0.9046	0.8585	0.7827	0.4062	0.4122	0.4243	0.5148	0.5157	0.5235	1.5840	1.1659	0.8798
$(80, 100]$	0.9615	0.8798	0.7652	0.9693	0.8980	0.8090	0.4292	0.4305	0.4574	0.5240	0.5243	0.5432	1.5884	1.1208	0.8425
$(100, \infty]$	1.0230	0.9311	0.7915	1.0336	0.9525	0.8499	0.5146	0.5086	0.5512	0.5866	0.5856	0.6183	1.6343	1.1537	0.8691

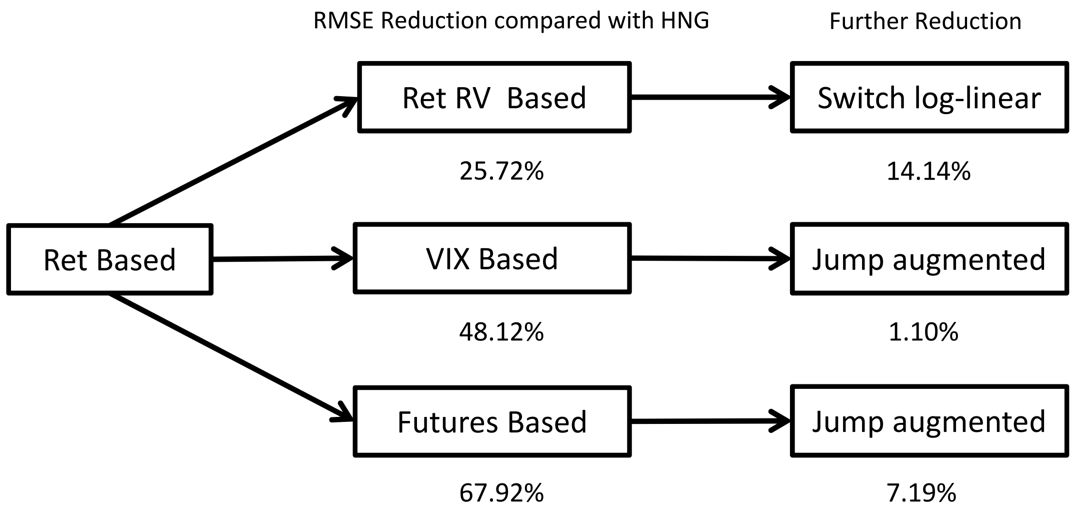
Note: This table reports the out-of-sample option pricing performance for each model. We evaluate the model's VIX option pricing ability through the root of mean square errors of option prices (RMSE). 1. We estimate the model by using the sample from 2006 to 2015 (the first ten years), and then use the estimated parameters to price the VIX options from 2016 to 2020. We summarize the results by option moneyness, and days to maturity. "Moneyness" is defined by the $m = \log F/K$, where F is the VIX futures price and K is the strike price. DTM denotes the number of days to maturity.

Table 6: Out-of-sample VIX Options Pricing Performance (Rolling Window)

	VIX-based Models						Futures-based Models						SPX-based Models		
	With Jump			No Jump			With Jump			No Jump					
Model	ARMA11	HAR(M)	HAR(Y)	ARMA11	HAR(M)	HAR(Y)	ARMA11	HAR(M)	HAR(Y)	ARMA11	HAR(M)	HAR(Y)	HNG	GARV	RG
Full RMSE	0.8136	0.7398	0.5654	0.8108	0.7490	0.5923	0.3600	0.3588	0.3585	0.4669	0.4668	0.4667	1.5316	1.0137	0.8538
Part 1: Evaluation By Moneyness m=log(F/K)															
$[-\infty,-0.4]$	0.3555	0.2697	0.2656	0.3142	0.3321	0.3669	0.2519	0.2547	0.2546	0.4508	0.4506	0.4505	0.5896	0.5857	0.3806
$(-0.4,-0.2]$	0.6125	0.5260	0.4683	0.5920	0.5565	0.5128	0.4039	0.3984	0.3980	0.5602	0.5597	0.5596	1.0904	0.9048	0.7417
$(-0.2,-0.1]$	0.7739	0.6963	0.5751	0.7636	0.7173	0.5851	0.4553	0.4498	0.4496	0.4884	0.4878	0.4878	1.5378	1.0577	0.9472
$(-0.1,0]$	0.9062	0.8212	0.6406	0.9029	0.8370	0.6460	0.4986	0.4955	0.4948	0.4428	0.4440	0.4438	1.8869	1.2206	1.0745
$(0,0.1]$	1.1064	1.0297	0.7703	1.1190	1.0203	0.7582	0.4909	0.4905	0.4902	0.4234	0.4232	0.4232	2.3226	1.4082	1.1849
$(0.1,\infty]$	1.4898	1.3972	0.9821	1.5157	1.3691	0.9666	0.3371	0.3373	0.3370	0.3943	0.3944	0.3945	2.7380	1.5861	1.4133
Part 2: Evaluation By Days to Maturity (DTM)															
$[0,20]$	0.4240	0.3335	0.3285	0.4180	0.3500	0.3429	0.3148	0.3030	0.3039	0.3818	0.3824	0.3827	1.1330	0.7641	0.7922
$(20,40]$	0.6144	0.5624	0.4527	0.6097	0.5692	0.4740	0.3525	0.3527	0.3497	0.4549	0.4528	0.4521	1.4030	0.9374	0.7687
$(40,60]$	0.9555	0.8818	0.7341	0.9612	0.8850	0.7528	0.3427	0.3516	0.3513	0.4861	0.4879	0.4877	1.7048	1.1999	0.9562
$(60,80]$	0.8951	0.8375	0.6214	0.8925	0.8441	0.6525	0.3308	0.3377	0.3413	0.4710	0.4716	0.4726	1.6434	1.1147	0.8515
$(80,100]$	0.9680	0.8782	0.6385	0.9650	0.8905	0.6774	0.3582	0.3596	0.3625	0.4798	0.4800	0.4812	1.6635	1.0570	0.8683
$(100,\infty]$	1.0242	0.9254	0.6532	1.0186	0.9391	0.6885	0.4462	0.4388	0.4356	0.5370	0.5362	0.5350	1.7267	1.0846	0.9399

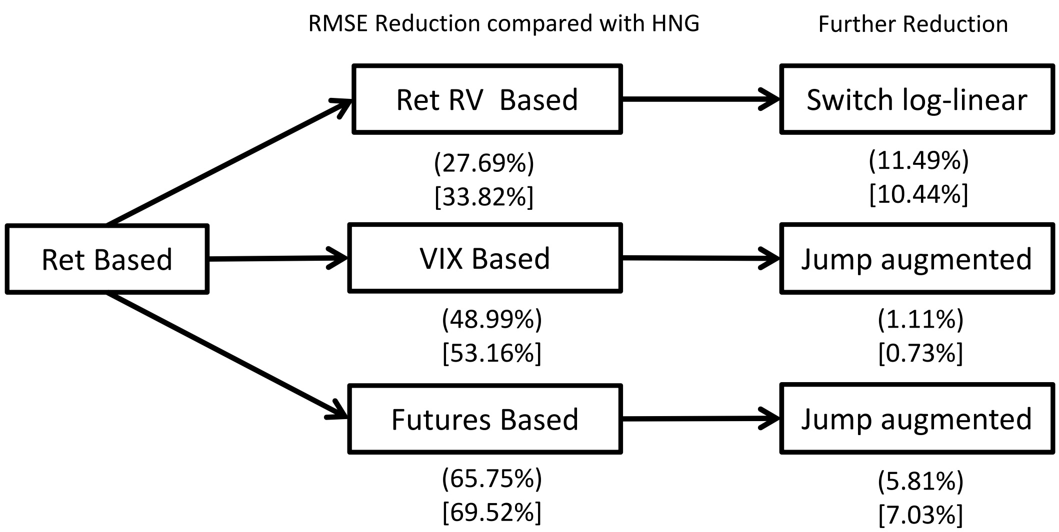
Note: This table reports the out-of-sample option pricing performance for each model. We evaluate the model's VIX option pricing ability through the root of mean square errors of option prices (RMSE). Our out-of-sample pricing analysis is based on a rolling window of ten years, with the parameters updated on a monthly basis. We evaluate the out-of-sample pricing errors from 2016 to 2020 (the observations in first ten years from 2006 to 2015 are used as a pre-sample to obtain the first set of parameters). We summarize the results by option moneyness, and days to maturity. "Moneyness" is defined by the $m = \log F/K$, where F is the VIX futures price and K is the strike price. DTM denotes the number of days to maturity.

Figure 1: In-sample RMSE reduction across different settings



Note: RMSE reduction is calculated against Heston-Nandi GARCH model (Ret Based) for the first layer (RMSE reduction compared with HNG) and the "Further reduction" is calculated with "additional" RMSE reduction (in terms of % reduction against Heston-Nandi GARCH model). For example, the Ret RV based model GARV reduces RMSE by 25.72% and a further 14.14% reduction can be achieved when we switch to log-linear RG model. If one directly compare RG with Heston-Nandi GARCH, the RMSE reduction is 39.81% (=25.72%+14.14%). The RMSE for VIX/Future based model and the corresponding jump augmented model is the simple average across ARMA(1,1), HAR(M) and HAR(Y) in each sub-categories.

Figure 2: Out-of-sample RMSE reduction across different settings



Note: RMSE reduction is calculated against Heston-Nandi GARCH model (Ret Based) for the first layer (RMSE reduction compared with HNG) and the "Further reduction" is calculated with "additional" RMSE reduction (in terms of % reduction against Heston-Nandi GARCH model). As two out-of-sample settings are used, we report "estimate-and-forget" results in parenthesis and "rolling-window" results in brackets. For example, under "estimate-and-forget" method, the Ret RV based model GARV reduces RMSE by 27.69% and a further 11.49% reduction can be achieved when we switch to log-linear RG model. If one directly compare RG with Heston-Nandi GARCH, the RMSE reduction is 39.18% (=27.69%+11.49%). Using "rolling-window", the corresponding numbers are 33.82%, 10.44% and 44.25%. The RMSE for VIX/Future based model and the corresponding jump augmented model is the simple average across ARMA(1,1), HAR(M) and HAR(Y) in each sub-categories.