Journal of Futures Markets

RESEARCH ARTICLE OPEN ACCESS

Pricing VXX Options With Observable Volatility Dynamics From High-Frequency VIX Index

Shan Lu

Kent Business School, University of Kent, Canterbury, UK

Correspondence: Shan Lu (s.lu@kent.ac.uk)

Received: 25 August 2024 | **Revised:** 7 January 2025 | **Accepted:** 7 April 2025

Funding: The author received no specific funding for this work.

Keywords: heterogeneous autoregressive model | realized variance | VIX exchange-traded product | VXX options

ABSTRACT

This paper develops a discrete-time joint analytical framework for pricing volatility index (VIX) and VXX options consistently. We show that our framework is more flexible than continuous-time VXX models as it allows the information contained in the high-frequency VIX index to be incorporated for the joint pricing of VIX and VXX options, and the joint pricing formula is derived. Our empirical analysis shows that the model that utilizes the realized variance (RV) computed from the high-frequency VIX index data significantly outperforms the model that does not rely on the VIX RV in the joint pricing both in-sample and out-of-sample, reinforcing the beliefs that high-frequency data are informative about the derivatives pricing

JEL Classification: C52, G13

1 | Introduction

This paper develops a discrete-time joint analytical framework of pricing volatility index (VIX) and VXX options consistently, in contrast to all the previous works on VXX option pricing, which are in the continuous-time setting. By specifying a discrete-time model for the VIX dynamics—the family of the heterogeneous autoregressive (HAR) generalized autoregressive conditionally heteroskedastic (GARCH) models, we derive the implied VXX dynamics and the VXX option pricing formula. The advantage of the discrete-time pricing framework is that it allows a variety of exogenous information to be incorporated for the pricing of VXX options, which prior continuous-time VXX models are incapable of. In particular, in this paper, we embed the realized variance (RV) of the high-frequency VIX index into the joint pricing, allowing us to exploit the information content of RV computed from the high-frequency VIX index and ask the question of whether high-frequency VIX index data contain valuable information about the joint pricing of VIX and VXX options.

Issued by Barclays in 2009, the VXX—the most actively traded exchange-traded note (ETN) on the VIX futures—is an unsecured debt obligation with no coupons, whose final redemption value at maturity depends on the value of the S&P 500 VIX short-term futures index total return (SPVXSTR). The SPVXSTR index tracks the performance of a portfolio that takes a long position in the nearest and second-nearest maturing VIX futures with a synthetic 30-day constant maturity. The VXX has since grown in popularity among investors to manage equity market volatility risk, and there are mainly two reasons: first, VIX derivatives are only accessible to large institutional investors due to their large notional sizes; in contrast, the VXX is more accessible to retail investors due to its small notional sizes, its high correlation with the VIX index makes it an affordable diversification tool for retail investors to manage equity market volatility risk. Second, some institutional investors, such as pension funds and mutual funds, are prohibited by regulations from taking positions in financial derivatives, which leaves them vulnerable to equity market volatility risk; the VXX offers these organizations a suitable hedging vehicle, though recent empirical studies have questioned the effectiveness of the

VIX futures ETNs as a diversification tool for equity holdings (see, e.g., Deng et al. 2012; Alexander and Korovilas 2013; Alexander et al. 2015, 2016). The top figure of Figure 1 exhibits the historical evolutions of the VXX/VXXB price, the VIX index, and the 30-day VIX futures price from January 30, 2009 to December 24, 2024. The bottom figure of Figure 1 shows the log price of VXX/VXXB and the reverse splits in the same period (VXX Price and Trade Volume 2024). VXX/VXXB price is adjusted for reverse splits, and the 30-day VIX futures price is obtained by linear interpolation of market VIX futures prices. Series A of VXX expired on the 10th anniversary of its launch on January 29, 2019, and series B of VXX (VXXB) was launched on January 18, 2018; both series are identical products, and prices of both series are almost identical during the overlapping period between January 2018 and January 2019 (see Lin and Zhang 2022, Figure 4).[1] During the period, VXX lost more than 99.99% of its value since inception due to the cost of rollover of VIX futures contracts, which prompted VXX to undergo eight reverse splits, with the most recent reverse split undertaken on July 24, 2024. In contrast, the VIX loses only about 68.18% in the same period. It is clear from the bottom figure that VXX and VIX are closely related: when the VIX spikes, VXX also experiences peaks. For example, VIX reached an all-time high of 82.69 on March 16, 2020 due to the impact of the outbreak of the Covid-19 pandemic; VXX reached a local peak of 4416 (reverse-split adjusted value) 2 days later on March 18, 2020. The top figure of Figure 2 shows the log trade volume of VXX/VXXB during the same period, which clearly shows the growing popularity of the VXX ETN.

VXX options were introduced in 2010 by the Chicago Board Options Exchange (CBOE). It is widely observed that the VXX options exhibit a positive skew due to the VXX's close relationship with the VIX index, making VXX options an ideal alternative to the VIX options to hedge against market downturns; by using VXX options, investors can also benefit from acquiring the underlying at a lower cost to cover their positions in the options market due to VXX's small notional sizes. The bottom figure of Figure 2 exhibits the daily trade volume of VXX options across all CBOE exchanges (including CBOE, BATS, C2, and EDGX) since its inception in May 2010; daily trade volume is the sum of trade volumes across all these exchanges (CBOE VXX Options Trade Volume 2024).[2] The figure demonstrates the increasing trading and popularity of VXX options.

This paper is motivated by two strands of literature. First, extended econometric literature documents that realized measures such as realized volatility and jumps computed from high-frequency data provide accurate estimates of the volatility process and improve the performance of forecasting future volatility. For example, Shephard and Sheppard (2010) found the HEAVY model based on realized volatility from high-frequency data provides better fits of the data than traditional GARCH models during the credit crunch. Hansen et al. (2012) developed a realized GARCH model that utilizes both returns and the realized volatility measure from high-frequency data and found that the realized GARCH model improves the volatility forecasts compared with standard GARCH models. Hansen and Huang (2016) further improved volatility forecasting performance by developing an exponential GARCH model that utilizes multiple realized volatility measures. Corsi (2009) developed the HAR model, which considers different realized volatility components over different time horizons, and found that the model greatly improved volatility forecasting performance. More recently, Bu et al. (2023) showed that a model that utilizes decomposed realized jump measures by activity and signs generates significantly better volatility forecasts.

Second, many studies apply realized measures such as realized volatility and realized jump variations computed from high-frequency data in option valuation and show that realized volatility and jumps contain valuable information about the pricing of derivative products, as these realized measures allow for additional risk premiums linked to the asset volatility shocks. On the one hand, in terms of equity index option pricing, Corsi et al. (2013) and Alitab et al. (2020) utilized a HAR model and found that the realized volatility computed from the high-frequency S&P 500 index improves the pricing of the S&P 500 index options, while Christoffersen et al. (2014) come to the same findings under the generalized affine realized volatility (GARV) framework. Feunou and Okou (2019) further showed that separating the realized upside and downside volatilities of the high-frequency S&P 500 index further improves the pricing of S&P 500 index options. In addition, Christoffersen et al. (2015) and Z. Pan et al. (2021) demonstrated that the realized jump variations from the high-frequency S&P 500 index contain valuable information about the S&P 500 index option pricing. On the other hand, in terms of VIX derivatives pricing, Q. Wang and Wang (2021) and Tong and Huang (2021) found that realized volatility computed from high-frequency S&P 500 improves VIX futures and option pricing, respectively, under the affine realized volatility model. By directly modeling the VIX index using the family of HAR-GARCH models, Jiang et al. (2022) find that the realized volatility and observable dynamic jumps computed from the high-frequency VIX index improve VIX futures pricing, whereas Qiao and Jiang (2023) find that realized semivariances of the high-frequency VIX index further improve VIX futures pricing; Guo et al. (2024) show that separate modeling of realized upside and downside variances of the high-frequency VIX index improve the model's performance in pricing both VIX futures and options.

In terms of VXX option pricing models, there is only limited and scarce literature; however, none of the VXX option pricing models in the literature allow information contained in the high-frequency VIX or VXX data to be exploited for VXX option pricing. By directly modeling the VXX dynamics, Bao et al. (2012), Tan et al. (2021), and Cao et al. (2021) provided semiclosed-form VXX option pricing formulas based on logarithmic stochastic volatility models and examined the roles of VXX jump structure and default risk in pricing VXX options. Gehricke and Zhang (2020a) directly fitted quadratic polynomial functions of the VXX implied volatility and price VXX options by polynomial regressions of the VXX implied volatility. In contrast, by modeling the VIX index instead of modeling the VXX directly, Grasselli and Wagalath (2020) and Lin and Zhang (2022) developed joint pricing frameworks for consistently pricing VIX and VXX options based on the logarithmic Ornstein–Uhlenbeck model with stochastic volatility (LOUSV). Lu (2023) extended the joint pricing framework by adding self-exciting jumps and showed that VIX jump clustering is

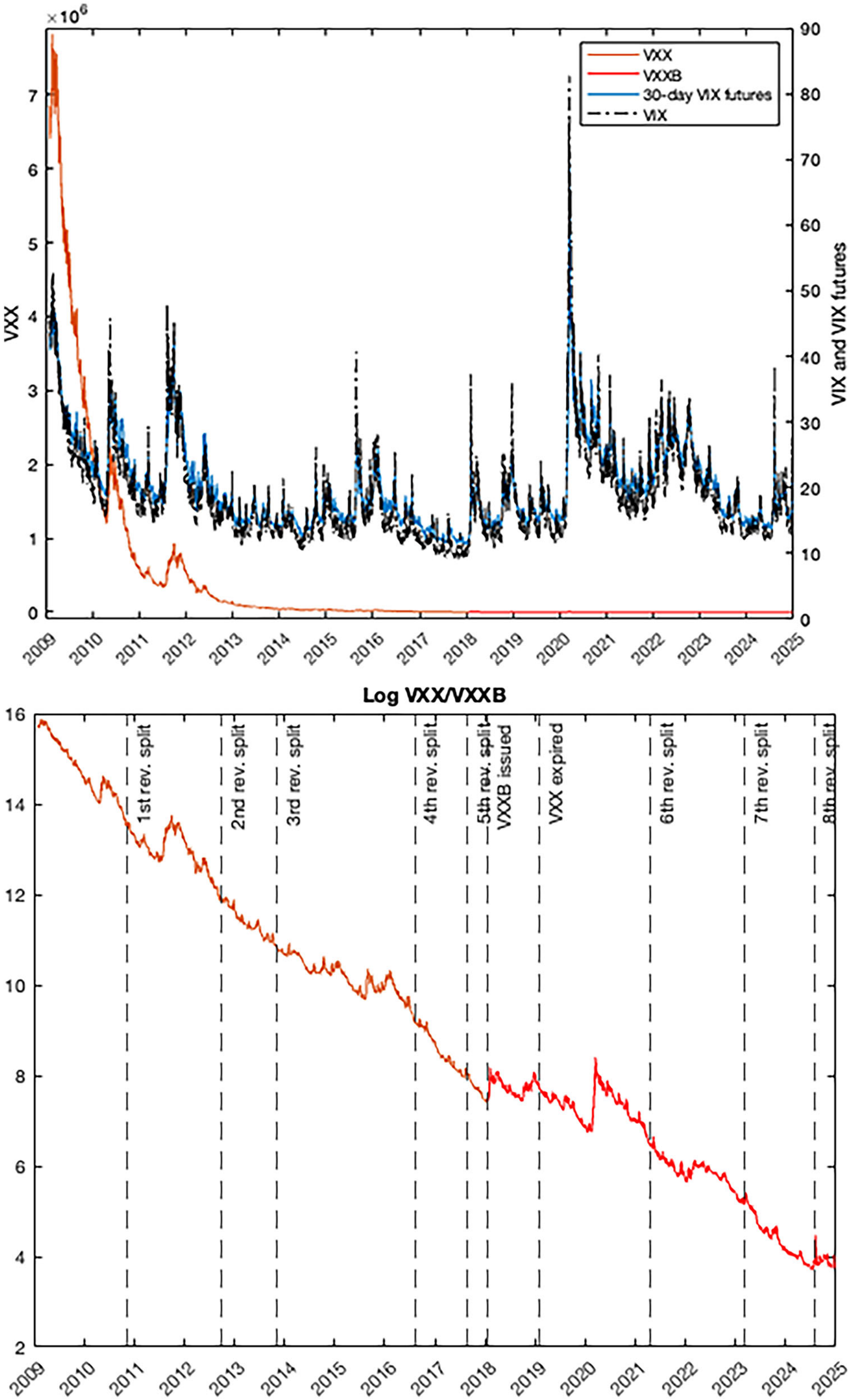

FIGURE 1 | VXX/VXXB price, VIX index, and 30-day VIX futures price. The top figure exhibits the daily prices of VXX/VXXB, VIX index, and 30-day VIX futures from January 30, 2009 to December 24, 2024. The bottom figure exhibits the log price of VXX/VXXB and the reverse splits that VXX/VXXB has undergone during the same period. The VXX/VXXB price is adjusted for reverse splits. The 30-day VIX futures price is obtained by linear interpolation of market VIX futures prices. "rev. split" stands for the reverse split. VIX, volatility index. [Color figure can be viewed at wileyonlinelibrary.com]

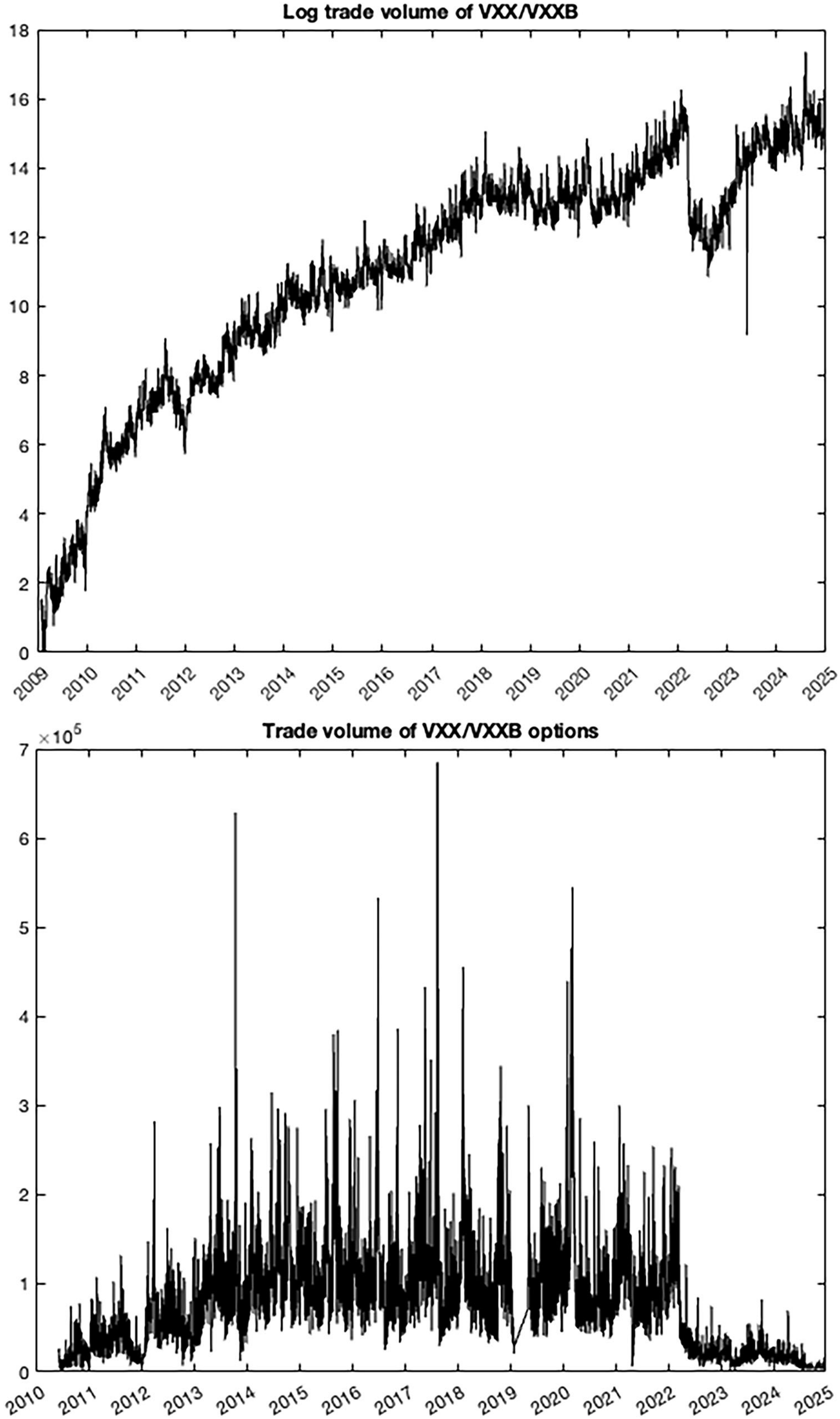

FIGURE 2 | Trade volumes of VXX/VXXB and options on VXX/VXXB. The top figure exhibits the daily log trade volume of VXX/VXXB from January 30, 2009 to December 24, 2024. The bottom figure exhibits the daily trade volume of options written on VXX/VXXB across CBOE and its subsidiary exchanges (BATS, C2, and EDGX) from May 28, 2010 to December 24, 2024. The trade volume of VXX/VXXB options is the sum of trade volumes across all CBOE exchanges. CBOE, Chicago Board Options Exchange.

important for the joint pricing of VIX and VXX options. Grasselli et al. (2023) developed a joint pricing framework based on the stochastic local volatility model, but their approach does not allow for a full model calibration to be carried out.

There are several reasons for our discrete-time joint pricing framework to adopt the consistent pricing approach by modeling the VIX index directly. First, a joint pricing framework allows for the theoretical linking of the characteristics of the VIX and its derivatives to those of its ETNs and ETN options in a tractable manner. Second, a joint pricing model also provides a tool to study and predict the behavior of the two markets consistently. Third, although unified pricing of SPX, VIX, and VXX and their derivatives remains a problem to be solved, developing a joint pricing framework of VIX and VXX and their derivatives is a step toward such a unified pricing model.

This paper makes several contributions to the literature, and we summarize them below:

- To the best of our knowledge, this paper is the first to develop a discrete-time pricing framework for consistently pricing VIX and VXX options. More importantly, we show that the pricing framework is more flexible and allows one to incorporate exogenous information, compared with previous VXX option pricing models in the literature, which are all in the continuous-time setting and are incapable of incorporating exogenous information for VXX option pricing.
- Second, we derive the formula for joint pricing VIX and VXX options that utilize the information contained in the RV from the high-frequency VIX data, which has never been done in the literature on pricing VXX options and, more generally, options on VIX futures ETNs.
- Third, we provide the first empirical work that exploits the information in the RV computed from the high-frequency VIX index for the joint VIX and VXX option pricing.

By fitting our model to the real market VIX and VXX derivatives data, we show that the model that utilizes information contained in the RV computed from the high-frequency VIX index significantly outperforms the model that does not rely on the realized measure from the high-frequency VIX, indicating that high-frequency VIX index data contain valuable information about the joint pricing of VIX and VXX options. Our findings contribute to the literature on applying realized measures computed from high-frequency data in option valuation and reinforce the beliefs that high-frequency data are informative about derivatives pricing.

Our paper is also related to the literature that studies the dynamics of the VXX and the VXX options market. Eraker and Wu (2017) provide theoretical foundations to explain the negative risk premia of the VXX by using an equilibrium model of the equity dynamics; Gehricke and Zhang (2018, 2020b) establish the theoretical link between the S&P 500 index, the VIX, and the VXX based on a continuous-time stochastic model of the stock price; Yoon et al. (2022) study the relationship among the implied volatility of equity, VIX, and VXX options.

The rest of the paper unfolds in the following way: Section 2 introduces the specifications of the discrete-time VIX models in the literature that we employ as the base model. Section 3 provides the derived implied VXX return dynamics and closed-form VXX option pricing formulas based on the discrete-time VIX models. Section 4 provides details of the empirical analysis to assess the models' performance in joint pricing VIX and VXX options, including data, model estimation, and performance measures. Empirical results are discussed in Section 5. Section 6 concludes the paper.

2 | VIX Model

In this section, we set out the specifications of the discrete-time VIX models that we employ as the base model.

2.1 | Brief Review of VIX Models

There are two approaches to pricing VIX futures and options in the literature. The first approach is to specify a model for the SPX index and derive the implied VIX dynamics and VIX derivatives pricing formula based on the assumed equity dynamics in a consistent way, but this strand of literature is less relevant to the consistent pricing of VIX and VXX options.

The second approach is to directly specify a model for the VIX dynamics and derive the VIX derivatives pricing formula based on the assumed model of the VIX index. Stand-alone VIX models can be broadly divided into two categories: continuous-time and discrete-time models. Continuous-time VIX models include affine and nonaffine diffusion models of Goard and Mazur (2013), Mencia and Sentana (2013), Park (2016), Luo et al. (2019), Jing et al. (2020), and Yuan (2022). Discrete-time VIX models mainly include the family of HAR-GARCH models developed by Yin et al. (2021), Jiang et al. (2022), Tong et al. (2022), Qiao and Jiang (2023), and Guo et al. (2024).

On the basis of the continuous-time stand-alone VIX models, Grasselli and Wagalath (2020) and Lin and Zhang (2022) developed the framework of joint pricing VIX and VXX options using the LOUSV, while Lu (2023) further extended the joint pricing framework using the LOUSV model with Hawkes-type jumps.

This paper, in contrast, develops a consistent framework for pricing VIX and VXX options based on the discrete-time stand-alone VIX models: the family of HAR-GARCH models. We will work with the HAR-RV-GARCH and HAR-GARCH models in the rest of the paper. Our approach can be easily applied to derive consistent pricing formulas based on other models in the HAR-GARCH family.

2.2 | VIX Model Specifications

2.2.1 | HAR-RV-GARCH

Let $y_t = \ln(\mathrm{VIX}_t)$. Under the risk-neutral measure $\mathbb{Q}$, we assume the VIX dynamics follows a *HAR-RV-GARCH* model proposed by Jiang et al. (2022) and Qiao and Jiang (2023). Following their model specification,

$$y_t = \beta_0 + \sum_{i=1}^{p} \beta_i y_{t-i} + z_t, \tag{1}$$

$$h_{z,t} = \omega + b h_{z,t-1} + a \mathrm{RV}_t, \tag{2}$$

where $z_t = \sqrt{h_{z,t-1}} \varepsilon_{1,t}$, where $\varepsilon_{1,t} \sim N(0, 1)$. $h_{z,t}$ is the conditional variance of the log VIX at t, which is associated with the RV of the VIX index at t, RV_t. In Equation (2), RV_t is calculated using high-frequency VIX data. While Jiang et al. (2022) and Qiao and Jiang (2023) only allow the long memory of the VIX information up to a month ($p = 22$), we allow an arbitrary number of lags of logarithmic VIX in Equation (1).

Following Jiang et al. (2022) and Qiao and Jiang (2023), by introducing the measurement error $\varepsilon_{2,t} \sim N(0, 1)$, whose correlation with $\varepsilon_{1,t}$ is ρ, the measurement equation for the RV can be defined as

$$\begin{aligned} \mathrm{RV}_t = {} & h_{z,t-1} + \sigma((\gamma^*)^2 - \gamma^2) h_{z,t-1} \\ & + \sigma\left[(\varepsilon_{2,t} - \gamma^* \sqrt{h_{z,t-1}})^2 - (1 + (\gamma^*)^2 h_{z,t-1})\right], \end{aligned} \tag{3}$$

where γ and γ^* measure the leverage effects under the physical measure and the risk-neutral measure, respectively. To ensure the positivity and stationarity of the conditional variance dynamics, we impose the following conditions: $\omega, b, a, \sigma, \gamma > 0, \gamma^* > \gamma$, and $b + a + a\sigma((\gamma^*)^2 - \gamma^2) < 1$. In the estimation, we estimate γ and $\delta = \gamma^* - \gamma > 0$. The persistency of the conditional variance of the VIX is defined as $b + a + a\sigma((\gamma^*)^2 - \gamma^2)$, and the long-term variance is calculated as $\omega / [1 - (b + a + a\sigma((\gamma^*)^2 - \gamma^2))]$. Other discrete-time VIX models proposed in the literature that adopt the same settings of the conditional variance Equation in (2) and the measurement Equation in (3) include the HAR-DJI-GARCH model in Jiang et al. (2022), the HAR-RV-ud-GARCH model in Qiao and Jiang (2023), and the HAR-RSV model in Guo et al. (2024), and the measurement Equation in (3) is also consistent with the setup of the measurement equation in the GARV model in Christoffersen et al. (2014).

2.2.2 | HAR-GARCH

As an alternative model, we assume that the VIX index follows a *HAR-GARCH* model without the information from the realized volatility computed from the high-frequency VIX data. The HAR-GARCH model of the VIX has the following specification under $\mathbb{Q}$ (see Jiang et al. 2022; Qiao and Jiang 2023):

$$y_t = \beta_0 + \sum_{i=1}^{p} \beta_i y_{t-i} + z_t, \tag{4}$$

$$h_{z,t} = \omega + b h_{z,t-1} + a(\varepsilon_{t-1} - \gamma^* \sqrt{h_{z,t-1}})^2, \tag{5}$$

where we also allow an arbitrage number of lags of the VIX; $z_t = \sqrt{h_{z,t}} \varepsilon_t$, $\varepsilon_{1,t} \sim N(0, 1)$. $h_{z,t}$ is the conditional variance of the VIX. To ensure the positivity and stationarity of the conditional variance process, we impose the following parameter restrictions: $a > 0, \omega + a > 0$, and $b + a(\gamma^*)^2 < 1$. The persistency of the conditional variance of the VIX is $b + a(\gamma^*)^2$, and the long-term variance is $\omega / [1 - (b + a(\gamma^*)^2)]$.

Due to the arbitrary number of lags of the VIX, the aforementioned two models subsume several models in the literature: if we impose the constraints that $\beta_1 = \beta_d, \beta_i = \beta_w / 4$ when $2 \le i \le 5, \beta_i = \beta_m / 17$ when $6 \le i \le 22$, the model reduces to the HAR-GARCH (M) model similar to that in Yin et al. (2021), Jiang et al. (2022), and Qiao and Jiang (2023), which accounts for long memory of VIX up to a month. If we add two additional restrictions together with the restrictions above: $\beta_i = \beta_q / 41$ when $23 \le i \le 63, \beta_i = \beta_Y / 189$ when $64 \le i \le 252$, the model becomes the HAR-GARCH (Y) model incorporating past VIX information up to a year, similar to that in Tong et al. (2022) but with time-varying volatility. In addition, as mentioned by Yin et al. (2021), the following models can also obtained by imposing the following restrictions: the mixed data sampling model when $\beta_i = \left(\left(\frac{1}{p}\right)^{\theta_0 - 1}\left(1 - \frac{1}{p}\right)^{\theta_1 - 1}\right)\left(\sum_{j=1}^{p}\left(\frac{j}{p}\right)^{\theta_0 - 1}\left(1 - \frac{j}{p}\right)^{\theta_1 - 1}\right)^{-1}$, and the exponentially weighted moving average model when $\beta_i = \theta^{p-i}(1 - \theta)(1 - \theta^p)^{-1}$.

2.3 | Pricing VIX Futures and Options

2.3.1 | HAR-RV-GARCH

By exploring the results of Yin et al. (2021), Jiang et al. (2022), and Qiao and Jiang (2023) in the context of HAR-GARCH models for VIX futures pricing and those of Tong et al. (2022) and Guo et al. (2024) in the context of HAR(-GARCH) models for VIX option pricing, under the model in (1)–(3), the moment generating function (MGF) of y_T at time t is given by

$$\begin{aligned} f^{\mathrm{VIX}}(\phi; t, T) = E_t^{\mathbb{Q}}[\mathrm{e}^{\phi y_T}] = \exp\Bigg(& A(\phi; t, T) \\ & + B(\phi; t, T) h_{z,t} + \sum_{i=1}^{p} D_i(\phi; t, T) y_{t+1-i} \Bigg). \end{aligned} \tag{6}$$

The coefficients of the above MGF satisfy the following recursive relationships:

$$\begin{aligned} A(\phi; t, T) = {} & A(\phi; t+1, T) + B(\phi; t+1, T)(\omega - a\sigma) \\ & + D_1(\phi; t+1, T)\beta_0 \\ & - \tfrac{1}{2}\ln(1 - 2B(\phi; t+1, T)a\sigma), \end{aligned} \tag{7}$$

$$\begin{aligned} B(\phi; t, T) = {} & B(\phi; t+1, T)[b + a + a\sigma((\gamma^*)^2 - \gamma^2)] \\ & \frac{1}{2}(D_1(\phi; t+1, T))^2 - B(\phi; t+1, T)a\sigma \\ & + \frac{\begin{array}{c}(D_1(\phi; t+1, T))^2(1 - \rho^2) \\ + 2B(\phi; t+1, T)a\sigma\gamma^*(B(\phi; t+1, T)a\sigma\gamma^* \\ - D_1(\phi; t+1, T)\rho)\end{array}}{1 - 2B(\phi; t+1, T)a\sigma}, \end{aligned} \tag{8}$$

and

$$D_i(\phi;t,T)=\begin{cases}D_1(\phi;t+1,T)\beta_i+D_{i+1}(\phi;t+1,T), & 1\le i<p,\\ D_1(\phi;t+1,T)\beta_i, & i=p\end{cases} \quad (9)$$

subject to the following terminal conditions: $A(\phi;T,T)=B(\phi;T,T)=0$, and $D_1(\phi;T,T)=\phi$, and $D_i(\phi;T,T)=0$ for $2\le i\le p$.

2.3.2 | HAR-GARCH

Similarly, by exploring the results of Jiang et al. (2022), Qiao and Jiang (2023), and Tong et al. (2022), given the VIX dynamics in (4) and (5), the MGF of y_T at time t is

$$f^{\text{VIX}}(\phi;t,T)=\exp\left(A(\phi;t,T)+B(\phi;t,T)h_{z,t+1}+\sum_{i=1}^{p}D_i(\phi;t,T)y_{t+1-i}\right) \quad (10)$$

with

$$\begin{aligned}A(\phi;t,T)&=A(\phi;t+1,T)+B(\phi;t+1,T)\omega+D_1(\phi;t+1,T)\beta_0-\frac{1}{2}\ln(1-2B(\phi;t+1,T)a),\\ B(\phi;t,T)&=B(\phi;t+1,T)b+\frac{\frac{1}{2}(D_1(\phi;t+1,T))^2+B(\phi;t+1,T)a(\gamma^*)^2-2a\gamma^*B(\phi;t+1,T)D_1(\phi;t+1,T)}{1-2B(\phi;t+1,T)a},\end{aligned}$$

and

$$D_i(\phi;t,T)=\begin{cases}D_1(\phi;t+1,T)\beta_i+D_{i+1}(\phi;t+1,T), & 1\le i<p,\\ D_1(\phi;t+1,T)\beta_i, & i=p\end{cases}$$

subject to the following terminal conditions: $A(\phi;T,T)=B(\phi;T,T)=0$, $D_1(\phi;T,T)=\phi$, and $D_i(\phi;T,T)=0$ for $2\le i\le p$.

2.3.3 | VIX Futures and Option Pricing

The time-t price of a VIX futures contract that matures on T is then given by

$$F_{t,T}=f^{\text{VIX}}(1;t,T) \quad (11)$$

and the time-t price of a VIX call option contract that matures on T with strike price K is given by

$$C^{\text{VIX}}(t,T,K)=\mathrm{e}^{-r(T-t)}[F_{t,T}\Pi_1-K\Pi_2] \quad (12)$$

with

$$\Pi_j=\frac{1}{2}+\frac{1}{\pi}\int_0^{\infty}\Re\left[\frac{e^{-i\phi\log(K)}\varphi_j(\phi;t,T)}{i\phi}\right]d\phi,$$

where $\varphi_1(\phi;t,T)=f^{\text{VIX}}(i\phi+1;t,T)/f^{\text{VIX}}(1;t,T)$, $\varphi_2(\phi;t,T)=f^{\text{VIX}}(i\phi;t,T)$; $\Re$ is the real part of a complex number, $i=\sqrt{-1}$; r is the interest rate. The integrals are calculated by using the Gauss–Laguerre quadrature of order 20. The VIX put option price is obtained by the put–call parity.

3 | VXX Option Pricing

In this section, first, we derive the implied VXX return dynamics in discrete time based on the assumed discrete-time VIX dynamics in Section 2 and show that the implied VXX return dynamics satisfy the martingale condition for a tradable asset. Second, we derive the VXX option pricing formula based on the implied VXX return dynamics in discrete time. The results in this section are the main contributions of this paper and are new to the literature on VXX option pricing.

3.1 | Implied VXX Return Dynamics

We consider the daily rebalanced VXX, where 30-day constant maturity VIX futures are rolled daily to construct the VXX. In this case, a 30-day constant maturity futures contract is held for 1 day and rolled to the second 30-day constant maturity contract the next day. As shown in Lin and Zhang (2022), the relationship between the current VXX and VXX τ-day period later is

$$\text{VXX}_{t+\tau}=\text{VXX}_t\cdot\frac{F_{t+1,t+30}}{F_{t,t+30}}\cdot\frac{F_{t+2,t+31}}{F_{t+1,t+31}}\cdot\frac{F_{t+3,t+32}}{F_{t+2,t+32}}\cdots\frac{F_{t+\tau,t+\tau+29}}{F_{t+\tau-1,t+\tau+29}}\cdot\mathrm{e}^{r\tau} \quad (13)$$

and they derive the option pricing formula for VXX options using a general continuous-time stochastic model of the VIX.

Similar to the above case, we can obtain the implied daily rebalanced VXX dynamics from the dynamics of VIX futures in (11) based on the aforementioned discrete-time HAR-GARCH models.

Let $\tau=T-t$, and since the values of $A(\phi;t,T)$, $B(\phi;t,T)$, and $D_s(\phi;t,T)$ do not depend on t, but on τ, we simplify the notation and use $A(\phi,\tau)$, $B(\phi,\tau)$, and $D_s(\phi,\tau)$ in the rest of the paper, instead.

3.1.1 | HAR-RV-GARCH

First, consider a 1-day period, from the end of day t to the end of day $t+1$, the 30-day VIX futures is rolled once, and we have the relationship between VXX_{t+1} and VXX_t:

$$\frac{\mathrm{VXX}_{t+1}}{\mathrm{VXX}_t} = \frac{F_{t+1,t+30}}{F_{t,t+30}} \cdot e^r = \exp\begin{pmatrix} r + A(1,29) - A(1,30) + B(1,29)h_{z,t+1} \\ - B(1,30)h_{z,t} \\ + \sum_{i=1}^{p} D_i(1,29)y_{t+2-i} - \sum_{i=1}^{p} D_i(1,30)y_{t+1-i} \end{pmatrix}.$$

Taking logarithm on both sides, we have the daily log-return dynamics for the VXX:

$$\ln\frac{\mathrm{VXX}_{t+1}}{\mathrm{VXX}_t} = R_{t+1}^{\mathrm{VXX}} = r + A(1,29) - A(1,30) + B(1,29)h_{z,t+1} - B(1,30)h_{z,t} + \sum_{i=1}^{p} D_i(1,29)y_{t+2-i} - \sum_{i=1}^{p} D_i(1,30)y_{t+1-i}, \quad (14)$$

$$h_{z,t} = \omega + bh_{z,t-1} + a\mathrm{RV}_t. \quad (15)$$

Proposition 1 (Implied daily VXX return dynamics is a martingale). *The daily logarithmic VXX return process R_{t+1}^{VXX} as specified in (14) and (15) under the risk-neutral measure $\mathbb{Q}$ is a martingale during a whole day, from the end of day t to the end of day $t+1$, as it can be shown that*

$$E_t^{\mathbb{Q}}\left[\exp\left(R_{t+1}^{\mathrm{VXX}}\right)\right] = e^r.$$

Proof. See Appendix A. □

The implied daily VXX return dynamics in (14) and (15) can be simplified to the following by using (7)–(9) and (15):

TABLE 1 | Summary statistics for VIX derivatives.

	2012 (In-sample) Time-to-maturity (days)				2013 (Out-of-sample) Time-to-maturity (days)			
	All	**≤ 30**	**30–90**	**> 90**	**All**	**≤ 30**	**30–90**	**> 90**
A: Futures contracts								
Number of contracts	1034	161	403	470	1120	142	385	593
Average prices	23.0545	19.3729	21.8888	25.3152	17.3278	15.3756	16.6259	18.2510
B: Number of option contracts								
All	4468	592	2186	1690	3403	332	1542	1529
Call	3078	462	1533	1083	2616	268	1223	1125
Put	1390	130	653	607	787	64	319	404
C: Mean statistics								
Average prices	1.4379	0.7640	1.3238	1.8215	0.9884	0.5084	0.8542	1.2279
Average IV	0.8886	1.1976	0.9422	0.7110	0.8110	1.0725	0.8874	0.6926
Average trade volume	3837	10,568	4138	1091	4987	19,875	5736	999
Average open interests	36,835	88,052	43,125	10,758	44,299	144,615	56,204	10,511
D: Average prices by moneyness								
$m \leq -0.2$	0.6688	0.3107	0.5371	0.7885	0.4428	0.1833	0.3481	0.4981
$-0.2 < m \leq -0.06$	1.6550	0.6714	1.5395	2.2067	1.0984	0.4350	0.8856	1.4353
$-0.06 < m \leq 0.06$	2.9064	1.5132	2.7286	3.6980	2.0064	0.9835	1.7763	2.4992
$0.06 < m \leq 0.2$	2.2834	1.1333	2.1521	2.9511	1.5549	0.6850	1.4068	1.9968
$0.2 < m \leq 0.4$	1.3546	0.5997	1.2902	1.9391	0.9557	0.3757	0.8673	1.2744
$0.4 < m \leq 0.6$	0.7971	0.3306	0.7221	1.1314	0.5756	0.2302	0.4869	0.7568
$m > 0.6$	0.3856	0.2270	0.3402	0.4761	0.3059	0.1875	0.2599	0.3450
E: Average IV by moneyness								
$m \leq -0.2$	0.6438	0.8620	0.7091	0.5827	0.5380	0.7754	0.5832	0.5094
$-0.2 < m \leq -0.06$	0.7391	0.8461	0.7851	0.6322	0.6235	0.7470	0.6445	0.5784
$-0.06 < m \leq 0.06$	0.8128	0.9752	0.8574	0.6901	0.7012	0.8405	0.7388	0.6285
$0.06 < m \leq 0.2$	0.8862	1.1428	0.9207	0.7305	0.7871	1.0008	0.8264	0.6755
$0.2 < m \leq 0.4$	0.9901	1.3119	1.0043	0.7627	0.8860	1.1884	0.9275	0.7252
$0.4 < m \leq 0.6$	1.0464	1.4458	1.0842	0.7981	0.9433	1.3151	1.0154	0.7775
$m > 0.6$	1.0427	1.5337	1.1421	0.8210	0.9441	1.4377	1.0976	0.8119

Note: This table reports the summary statistics for VIX futures and options from January 3, 2012 to October 10, 2013. Option moneyness is defined as the logarithm of the ratio between the strike price of the option and the underlying asset forward price.
Abbreviations: IV, implied volatility; VIX, volatility index.

$$R_{t+1}^{\mathrm{VXX}} = r + H(B(1,29)) + \Lambda(B(1,29), B(1,30))h_{z,t} + B(1,29)aRV_{t+1} + D_1(1,29)z_{t+1}, \tag{16}$$

$$h_{z,t} = \omega + bh_{z,t-1} + a\mathrm{RV}_t \tag{17}$$

with the following constants implied from the VIX model

$$H(B(1,29)) = B(1,29)a\sigma + \frac{1}{2}\ln(1 - 2B(1,29)a\sigma), \tag{18}$$

$$\Lambda(B(1,29), B(1,30)) = B(1,29)b - B(1,30). \tag{19}$$

For simplicity, we let $H = H(B(1,29))$ and $\Lambda = \Lambda(B(1,29), B(1,30))$ in the rest of the paper.

Next, consider a τ-day period where 30-day constant maturity VIX futures are rolled τ times; the τ-day period VXX log-return can be represented by

$$\ln\frac{\mathrm{VXX}_{t+\tau}}{\mathrm{VXX}_t} = R_{t+\tau}^{\mathrm{VXX}} = \exp\begin{pmatrix} r\tau + \tau H + \Lambda\sum_{j=1}^{\tau} h_{z,t+j} + D_1(1,29)\sum_{j=1}^{\tau} z_{t+j} \\ + B(1,29)a\sum_{j=1}^{\tau} RV_{t+j} \end{pmatrix},$$

which is not a martingale from day t to day $t+\tau$.

3.1.2 | HAR-GARCH

Following the derivation shown in Section 3.1.1, consider a 1-day period, the daily implied VXX return dynamics based on the HAR-GARCH model of the VIX in (4) and (5) is as follows:

$$R_{t+1}^{\mathrm{VXX}} = r + H(B(1,29)) + \Lambda(B(1,29), B(1,30))h_{z,t+1} + B(1,29)a(\varepsilon_{t+1} - \gamma^*\sqrt{h_{z,t+1}})^2 + D_1(1,29)z_{t+1}, \tag{20}$$

TABLE 2 | Summary statistics for VXX options.

	2012 (In-sample) Time-to-maturity (days)				2013 (Out-of-sample) Time-to-maturity (days)			
	All	**≤ 30**	**30–90**	**> 90**	**All**	**≤ 30**	**30–90**	**> 90**
A: Number of option contracts								
All	6216	490	2499	3227	5633	1392	1789	2452
Call	4388	358	1828	2202	4131	972	1383	1776
Put	1828	132	671	1025	1502	420	406	676
B: Mean statistics								
Average BS prices	4.7903	1.5877	3.0973	6.5877	1.0927	0.4766	0.7662	1.6807
Average IV	0.8396	0.8430	0.8517	0.8297	0.7508	0.7244	0.7818	0.7432
Average trade volume	596	2668	780	138	1108	1890	1725	214
Average open interests	5059	16,584	4391	3827	7004	8720	8252	5119
C: Average BS prices by moneyness								
$m \le -0.2$	2.8747	0.4553	1.3398	3.5581	0.7817	0.1376	0.3236	0.9284
$-0.2 < m \le -0.06$	5.9915	1.1887	3.8434	9.5460	1.1830	0.3262	0.8859	2.3928
$-0.06 < m \le 0.06$	8.7652	3.0367	6.7146	12.7885	1.7129	0.7798	1.8266	3.4610
$0.06 < m \le 0.2$	7.0597	1.7644	5.1807	11.1991	1.3116	0.4445	1.2993	2.9217
$0.2 < m \le 0.4$	4.7494	0.8766	2.9103	8.0716	1.1149	0.2418	0.6904	2.0691
$0.4 < m \le 0.6$	3.4993	0.6029	1.4875	5.9415	0.8453	0.1536	0.3443	1.3678
$m > 0.6$	2.6962	—	0.9055	3.3199	0.5436	—	0.2240	0.6639
D: Average IV by moneyness								
$m \le -0.2$	0.6989	0.6923	0.6717	0.7106	0.6169	0.6338	0.5617	0.6333
$-0.2 < m \le -0.06$	0.7187	0.6647	0.7029	0.7502	0.6103	0.5729	0.5939	0.6666
$-0.06 < m \le 0.06$	0.7623	0.7297	0.7585	0.7774	0.6492	0.6224	0.6596	0.6926
$0.06 < m \le 0.2$	0.8155	0.8467	0.8140	0.8047	0.7341	0.7502	0.7226	0.7158
$0.2 < m \le 0.4$	0.8762	0.9721	0.8829	0.8399	0.8066	0.9086	0.8029	0.7474
$0.4 < m \le 0.6$	0.9341	1.1005	0.9685	0.8837	0.8464	1.0615	0.8973	0.7825
$m > 0.6$	0.9711	—	1.0494	0.9438	0.8848	—	0.9996	0.8416

Note: This table reports the summary statistics for VXX options from January 3, 2012 to October 10, 2013. Option moneyness is defined as the logarithm of the ratio between the strike price of the option and the underlying asset forward price. The strike price and the underlying price are adjusted according to the reverse-split ratio for VXX options traded before October 5, 2012, to avoid artificial jumps in the underlying price due to the reverse split in the VXX. BS prices are calculated using the market-quoted implied volatility (IV) and the Black–Scholes (BS) model.

$$h_{z,t} = \omega + bh_{z,t-1} + a(\varepsilon_{t-1} - \gamma^* \sqrt{h_{z,t-1}})^2 \quad (21)$$

with

$$H(B(1,29)) = \frac{1}{2}\ln(1 - 2B(1,29)a), \quad (22)$$

$$\Lambda(B(1,29), B(1,30)) = B(1,29)b - B(1,30). \quad (23)$$

And again, we let $H = H(B(1,29))$ and $\Lambda = \Lambda(B(1,29), B(1,30))$ in the rest of the paper.[3] It can be shown that the implied VXX return dynamics in (20) and (21) are a martingale during a whole day, and proofs are similar to those for the HAR-RV-GARCH model as shown in Appendix A and are thus omitted. Similarly, a τ-day period VXX return process is not a martingale from the end of day t to the end of day $t + \tau$.

The above results for the two models are analogous to those results for the VXX dynamics implied from continuous-time VIX models, as in Lin and Zhang (2022).

3.2 | Pricing VXX Options

To price VXX options, we derive the conditional MGF of VXX returns based on the HAR-RV-GARCH model, and we summarize this in the following proposition.

Proposition 2 (MGF of VXX return: HAR-RV-GARCH). *Given the implied VXX return dynamics in (16) and (17) under the risk-neutral measure $\mathbb{Q}$, the τ-step ahead MGF of VXX returns has the following exponential form:*

$$\Psi_{t,t+\tau}(u) = E_t^{\mathbb{Q}}\left[\exp\left(u\sum_{j=1}^{\tau} R_{t+j}^{VXX}\right)\right] = \exp(Q(u,\tau)h_{z,t} + R(u,\tau)) \quad (24)$$

with $Q(u,\tau)$ and $R(u,\tau)$ satisfying the following recursive relationships:

$$Q(u,\tau+1) = E(u, Q(u,\tau)),$$
$$R(u,\tau+1) = G(u, Q(u,\tau)) + R(u,\tau)$$

TABLE 3 | Summary statistics for the daily realized variance of the VIX index.

Year	Mean	Std. deviation	Skewness	Kurtosis	Minimum	Maximum	Jarque–Bera
2012 (In-sample)	0.0043	0.0058	5.4049	42.4528	2.9277e − 4	0.0594	1.7431e + 4
2013 (Out-of-sample)	0.0031	0.0030	2.3352	8.7307	3.7972e − 4	0.0169	446.3353

Note: This table reports the summary statistics for the daily realized variance of the VIX index computed from the high-frequency VIX data for the sample period from January 3, 2012 to October 10, 2013, including both the in-sample and out-of-sample. "Std. deviation" stands for the standard deviation, and "Jarque–Bera" stands for the Jarque–Bera test statistic.
Abbreviation: VIX, volatility index.

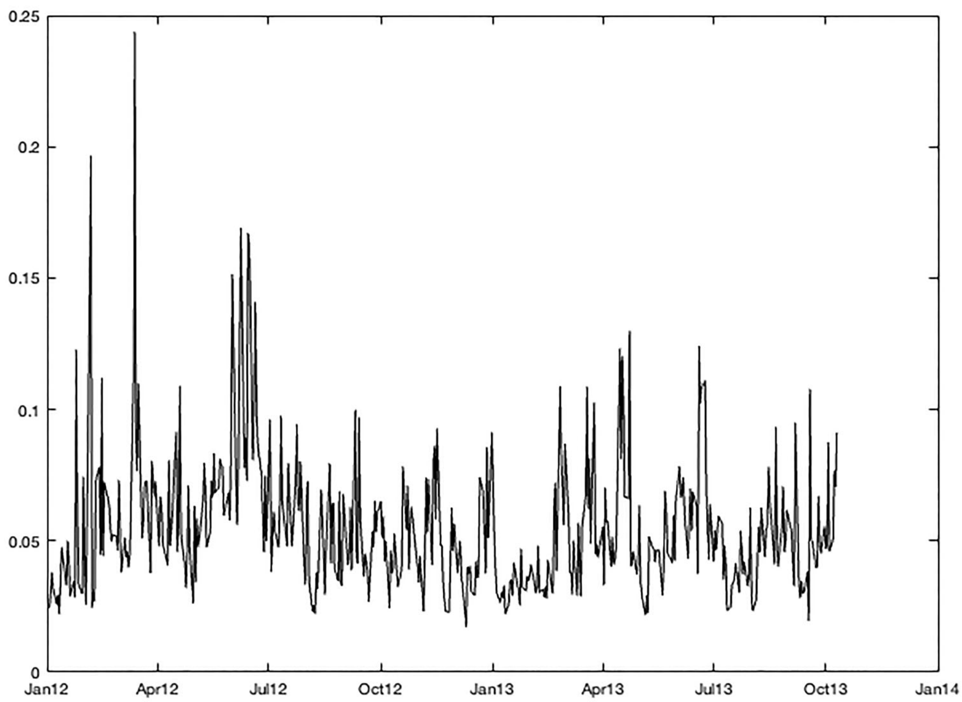

FIGURE 3 | Daily realized volatility of the VIX index. The figure exhibits the daily realized volatility of the VIX index calculated from 1-min grid of high-frequency VIX data for the sample period from January 3, 2012 to October 10, 2013. Realized volatility is the square root of realized variance, which is scaled to match the unconditional variance of the VIX index in the sample. VIX, volatility index.

subject to the initial conditions: $Q(u,1) = E(u,0)$, $R(u,1) = G(u,0)$, where

$$E(u,v) = u\Lambda + vb + (uB(1,29)a + va)(1 + \sigma((\gamma^*)^2 - \gamma^2)) + \frac{(uD_1(1,29)\rho - 2\sigma\gamma^*(uB(1,29)a + va))^2}{2(1 - 2\sigma(uB(1,29)a + va))} + \frac{1}{2}(uD_1(1,29))^2(1 - \rho^2),$$

$$G(u,v) = u(r + H) + v\omega - \sigma(uB(1,29)a + va) - \frac{1}{2}\ln(1 - 2\sigma(uB(1,29)a + va)).$$

Proof. See Appendix B. □

In the next proposition, we derive the MGF of VXX returns based on the HAR-GARCH model:

Proposition 3 (MGF of VXX return: HAR-GARCH). *Given the implied VXX return dynamics in (20) and (21) under the risk-neutral measure $\mathbb{Q}$, the τ-step ahead MGF of VXX returns has the following exponential form:*

$$\Psi_{t,t+\tau}(u) = E_t^{\mathbb{Q}}\left[\exp\left(u\sum_{j=1}^{\tau} R_{t+j}^{\text{VXX}}\right)\right] = \exp(Q(u,\tau)h_{z,t+1} + R(u,\tau)), \quad (25)$$

TABLE 4 | Parameter estimates.

	HAR-GARCH		HAR-RV-GARCH	
	Estimates	**Std. errors**	**Estimates**	**Std. errors**
β_0	0.0303	0.0197	0.0320	0.0012
β_d	0.9538	0.0169	0.9320	4.8556e − 4
β_w	9.5005e − 7	5.9825e − 8	1.3110e − 6	1.0330e − 6
β_m	0.0372	0.0239	0.0585	6.5447e − 4
ω	− 3.5841e − 4	5.8014e − 5	5.8450e − 4	2.1985e − 5
b	0.9035	0.0055	0.8915	0.0035
a	8.7173e − 4	7.1621e − 5	7.9373e − 5	3.2804e − 7
σ	—	—	2.1926	0.0124
γ	—	—	142.6691	1.1822
δ	—	—	0.1498	0.1927
γ^*	2.2089e − 6	8.0728e − 7	—	—
ρ	—	—	− 0.4128	0.3722
Persistency	0.9035		0.8990	
Average log-likelihood				
$\ln \mathcal{L}^{\text{F}}/N^{\text{F}}$	1.3992		1.4136	
$\ln \mathcal{L}^{\text{vix}}/N^{\text{vix}}$	− 0.7192		− 0.3428	
$\ln \mathcal{L}^{\text{vxx}}/N^{\text{vxx}}$	− 0.5780		− 0.0719	
Total	0.1020		0.9989	
Information criteria				
VIX futures				
AIC	− 2877.5		− 2901.4	
BIC	− 2838.0		− 2847.1	
VIX options				
AIC	6442.5		3085.4	
BIC	6493.7		3155.9	
VXX options				
AIC	7201.8		915.6	
BIC	7255.7		989.7	

Note: This table reports the parameter estimates by maximizing the joint log-likelihood function of VIX futures and options and VXX options pricing errors, where the pricing error is defined as the percentage error for the sample period from January 3 to December 31, 2012. The information criteria are calculated as AIC $= 2k - 2\ln\mathcal{L}$ and BIC $= k\ln N - 2\ln\mathcal{L}$, where k is the number of model parameters, N is the number of derivatives contracts, and $\ln\mathcal{L}$ is the value of the log-likelihood function of pricing errors. "Std. Errors" stands for standard errors of the parameter estimates.
Abbreviations: AIC, Akaike information criterion; BIC, Bayesian information criterion; GARCH, generalized autoregressive conditionally heteroskedastic; HAR, heterogeneous autoregressive; VIX, volatility index.

where the coefficients satisfy the following recursive relationships:

$$
\begin{aligned}
Q(u,\tau) &= u\Lambda + Q(u,\tau-1)b \\
&\quad + \frac{\frac{1}{2}(uD_1(1,29))^2 + \gamma^*(\gamma^* - 2uD_1(1,29))}{1 - 2[uB(1,29)a + Q(u,\tau-1)a]} \\
&\quad\quad\quad\quad\quad\quad \frac{+[uB(1,29)a + Q(u,\tau-1)a]}{}, \\
R(u,\tau) &= u(r+H) + Q(u,\tau-1)\omega + R(u,\tau-1) \\
&\quad - \frac{1}{2}\ln(1 - 2[uB(1,29)a + Q(u,\tau-1)a])
\end{aligned}
$$

subject to the following terminal conditions at time T: $Q(u,0) = R(u,0) = 0$.

Proof. See Appendix C. □

The time-t price of a VXX call option with strike price K and time-to-maturity τ can then be obtained by the Fourier inversion method:

$$C^{\mathrm{VXX}}(t,t+\tau,K) = \mathrm{VXX}_t P_1(t,t+\tau,K) - K\mathrm{e}^{-r\tau}P_2(t,t+\tau,K) \quad (26)$$

with

$$P_1(t,t+\tau,K) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \Re\left(\frac{\mathrm{e}^{-r\tau - iu\ln\left(\frac{K}{\mathrm{VXX}_t}\right)}\Psi_{t,t+\tau}(1+iu)}{iu}\right)du$$

$$P_2(t,t+\tau,K) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \Re\left(\frac{\mathrm{e}^{-iu\ln\left(\frac{K}{\mathrm{VXX}_t}\right)}\Psi_{t,t+\tau}(iu)}{iu}\right)du$$

where the integrals are also calculated by using the Gauss–Laguerre quadrature of order 20. The VXX put option price is obtained by the put–call parity.

4 | Empirical Analysis

In this section, we assess the performance of the HAR-RV-GARCH and HAR-GARCH models in terms of the joint pricing of VIX and VXX options. In particular, by comparing the performance of the two models, we are interested in the information content of the RV computed from the high-frequency VIX index data for the joint pricing.

It should be noted that there are no discrete-time models for VXX option pricing or joint pricing of VIX and VXX options proposed in the literature, and it is rare to compare discrete-time and continuous-time models in the option pricing literature. As a result, no VXX option pricing models in the literature are employed in the empirical part of the paper.

In our empirical analysis, both the HAR-RV-GARCH and HAR-GARCH models contain VIX observations realized over the period of up to 1 month (i.e., $p = 22$ in Equations 1 and 4), and we impose the constraints that $\beta_1 = \beta_d, \beta_i = \beta_w/4$

TABLE 5 | In-sample pricing performance: Overall.

	MAE			RMSE			MAPE		
	HAR-GARCH	**HAR-RV-GARCH**	**Reduction (%)**	**HAR-GARCH**	**HAR-RV-GARCH**	**Reduction (%)**	**HAR-GARCH**	**HAR-RV-GARCH**	**Reduction (%)**
VIX futures	**0.9766**	0.9777	−0.11	1.3202	**1.3094**	0.82	0.0443	**0.0437**	1.35
VIX options	0.4080	**0.3467**	15.02	0.5555	**0.4996**	10.06	0.3826	**0.2656**	30.58
VXX options	0.9042	**0.7476**	17.32	1.2333	**1.0581**	14.21	0.3137	**0.2038**	35.03

***t* statistics**

VIX futures	VIX options	VXX options
0.4103	**3.3184**	**5.2285**

Note: This table reports the overall in-sample pricing performance measured by the error metrics, including the Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE), and the Mean Absolute Percentage Error (MAPE). Pricing errors are computed by using parameter estimates in Table 4 and VIX and VXX data from January 3 to December 31, 2012. The t statistics are from the pairwise t test of Huang and Wu (2004) on the sample differences in the daily mean squared error. A positive t statistic that is larger than 1.645 indicates that the HAR-RV-GARCH model outperforms the HAR-GARCH model at a 5% significance level. Numbers in bold indicate the smallest error statistics or significance at a 5% significance level for the pairwise t test.

Abbreviations: GARCH, generalized autoregressive conditionally heteroskedastic; HAR, heterogeneous autoregressive; RV, realized variance; VIX, volatility index.

when $2 \leq i \leq 5, \beta_i = \beta_m/17$ when $6 \leq i \leq 22$, similar to the discrete-time VIX models in the literature.

4.1 | Data

We use an available data set containing data on daily end-of-day VIX futures and options and VXX options that were obtained from the CBOE and the 1-min high-frequency VIX index quotes that were obtained from Bloomberg and FirstRate Data (CBOE VIX Futures 2023; CBOE VIX Options 2023; CBOE VXX Options 2023, High-Frequency VIX Index 2023).[4] The sample spans from January 3, 2012 to October 10, 2013.[5] While we use daily VIX index and VIX futures with various maturities, to reduce the computational burden, we only use options data on Wednesdays in the sample, following Du and Luo (2019), Christoffersen et al. (2009), Eraker (2004), J. Pan (2002), and Chernov and Ghysels (2000). Using Wednesday options data is also due to the fact that Wednesday is the day of the week that is least likely to be a public holiday and is least affected by the day-of-the-week effect (see, e.g., Dumas et al. 1998; Augustyniak et al. 2025). Data from the year 2012 are used as the in-sample, and data from the year 2013 are used as the out-of-sample.

Commonly used filters are applied to options data with the intention to retain only liquid options: (1) In-the-money (ITM) options are removed when the option moneyness is defined as the logarithm of the ratio between the option strike price and the forward price of the underlying, ITM call options are call options with negative moneyness, and ITM put options are put options with positive moneyness; (2) options with a zero trade volume and bid prices smaller than \$0.1 are removed; (3) options with time-to-maturities smaller than 7 days and longer than 1 year are removed. For VIX options, the midpoint of the

TABLE 6 | In-sample pricing performance by moneyness and time-to-maturity.

	MAE		RMSE		MAPE	
	HAR-GARCH	HAR-RV-GARCH	HAR-GARCH	HAR-RV-GARCH	HAR-GARCH	HAR-RV-GARCH
VIX futures						
$\tau \leq 30$	**0.7659**	0.8062	**0.9466**	1.0050	**0.0406**	0.0425
$30 < \tau \leq 90$	**1.0474**	1.0921	**1.3699**	1.4057	**0.0493**	0.0502
$\tau > 90$	0.9882	**0.9383**	1.3850	**1.3159**	0.0413	**0.0386**
VIX options						
$m \leq -0.2$	0.2345	**0.1591**	0.3007	**0.2189**	0.5080	**0.2724**
$-0.2 < m \leq -0.06$	**0.2814**	0.3274	**0.3741**	0.4178	**0.2070**	0.2211
$-0.06 < m \leq 0.06$	**0.6236**	0.6372	0.8336	**0.8096**	**0.2311**	0.2334
$0.06 < m \leq 0.2$	0.6323	**0.5805**	0.7895	**0.7224**	0.2994	**0.2670**
$0.2 < m \leq 0.4$	0.4655	**0.3685**	0.5778	**0.4780**	0.3699	**0.2839**
$0.4 < m \leq 0.6$	0.3458	**0.2143**	0.4018	**0.2750**	0.4799	**0.2852**
$m > 0.6$	0.2073	**0.1025**	0.2332	**0.1385**	0.5777	**0.2823**
$\tau \leq 30$	0.2983	**0.2353**	0.3819	**0.3313**	0.4936	**0.3273**
$30 < \tau \leq 90$	0.4096	**0.3458**	0.5395	**0.4944**	0.4057	**0.2676**
$\tau > 90$	0.4444	**0.3868**	0.6231	**0.5525**	0.3138	**0.2414**
VXX options						
$m \leq -0.2$	0.5918	**0.5707**	0.8805	**0.8099**	0.3973	**0.2403**
$-0.2 < m \leq -0.06$	**0.5924**	0.7830	**0.8768**	1.0896	0.1693	**0.1596**
$-0.06 < m \leq 0.06$	**0.6593**	0.8828	**0.9340**	1.2231	**0.0827**	0.1010
$0.06 < m \leq 0.2$	**0.8738**	0.8926	**1.1828**	1.2393	0.1525	**0.1364**
$0.2 < m \leq 0.4$	1.0152	**0.7815**	1.3513	**1.1148**	0.2709	**0.1899**
$0.4 < m \leq 0.6$	1.1463	**0.7181**	1.4662	**0.9985**	0.4144	**0.2620**
$m > 0.6$	1.1886	**0.6707**	1.4812	**0.9456**	0.5503	**0.2774**
$\tau \leq 30$	0.4678	**0.3486**	0.6233	**0.4947**	0.3909	**0.2477**
$30 < \tau \leq 90$	0.6937	**0.5694**	0.9067	**0.8100**	0.3565	**0.2307**
$\tau > 90$	1.1335	**0.9462**	1.4947	**1.2694**	0.2689	**0.1763**

Note: This table reports the in-sample pricing performance by option moneyness and time-to-maturity. Pricing errors are computed by using parameter estimates in Table 4 and VIX and VXX data from January 3 to December 31, 2012. Error metrics, including the Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE), and the Mean Absolute Percentage Error (MAPE), are used. Option moneyness m is defined as the logarithmic of the ratio between the strike price and the underlying forward price. Numbers in bold indicate the smallest error statistics. τ denotes the time-to-maturity (days).
Abbreviations: GARCH, generalized autoregressive conditionally heteroskedastic; HAR, heterogeneous autoregressive; RV, realized variance; VIX, volatility index.

bid–ask price is used as the option price; for VXX options, as they are American style, we compute the European option price of VXX options by using the market-quoted implied volatility and the Black–Scholes model (see Carr and Wu 2010; Grasselli and Wagalath 2020; Cao et al. 2021; Lin and Zhang 2022; Lu 2023).[6] Since there was a reverse split on the VXX on October 5, 2012, we adjusted the underlying VXX value and the option strike price according to the reverse-split ratio (one-for-four) for VXX options traded before this date to avoid artificial jumps in the underlying VXX.

Tables 1 and 2 present the summary statistics for VIX derivatives and VXX options for both the in-sample and out-of-sample, respectively. Some patterns are identified: first, prices of both VIX and VXX options increase with time-to-maturity, and prices are the highest for near-the-money (NTM) options and

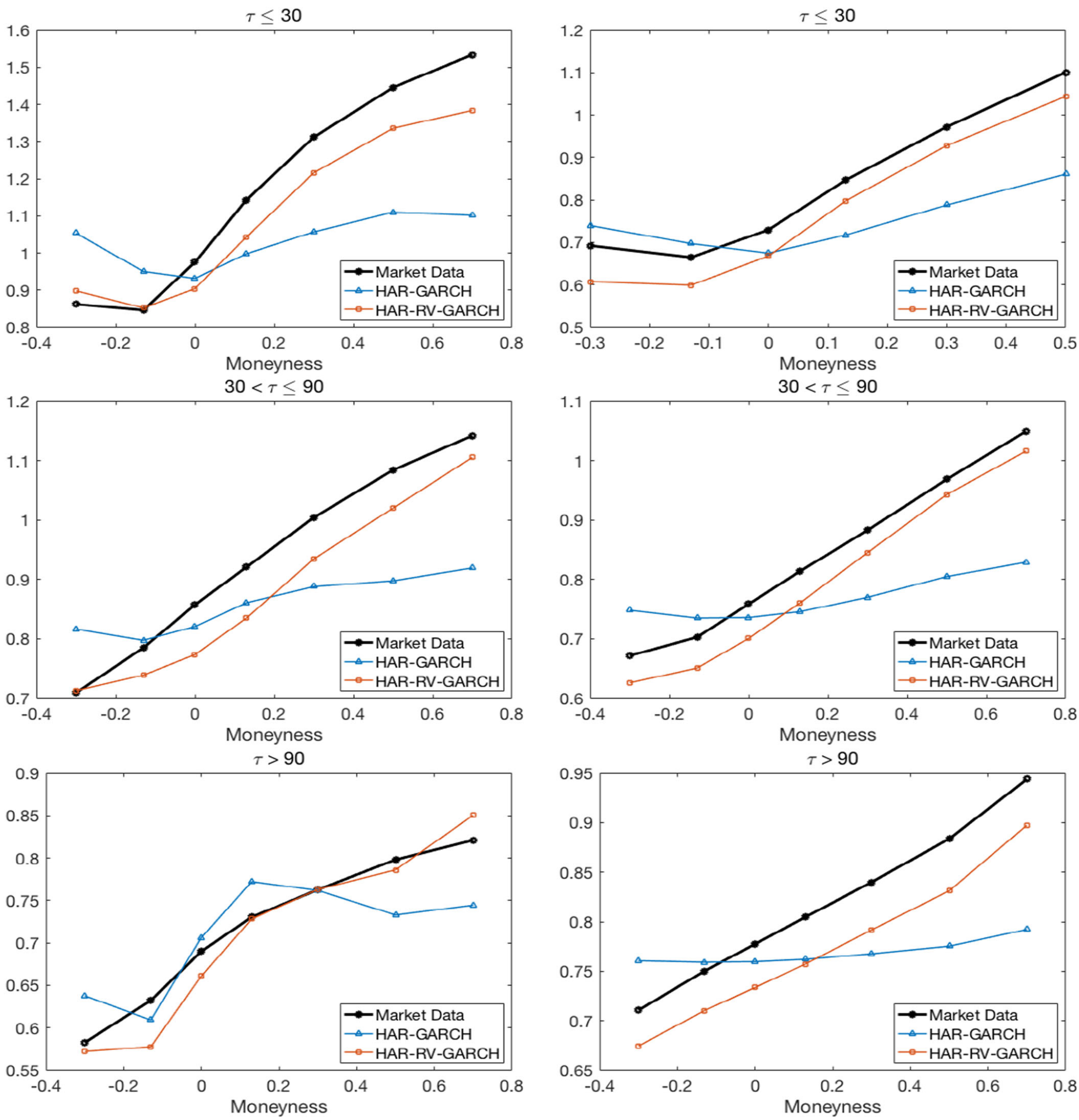

FIGURE 4 | In-sample fit to the implied volatility. The figure exhibits the average in-sample fit to the VIX and VXX implied volatilities across moneyness and time-to-maturity for the sample period from January 3 to December 31, 2012. Figures in the left column show the average in-sample fit to the VIX implied volatility, and figures in the right column show the average in-sample fit to the VXX implied volatility. Option moneyness is defined as the logarithm of the ratio between the strike price of the option and the underlying asset forward price. GARCH, generalized autoregressive conditionally heteroskedastic; HAR, heterogeneous autoregressive; RV, realized variance; VIX, volatility index. [Color figure can be viewed at wileyonlinelibrary.com]

decrease when moving deep into the money. Second, both VIX and VXX options exhibit positive skews; that is, the implied volatility increases with moneyness. Third, in both VIX and VXX options markets, there are more call options than put options, indicating both options are primarily used to hedge against a hike in the equity market volatility. Fourth, the average implied volatility of VXX options is lower than that of VIX options. In addition, the average VXX option price during 2013 was lower than that in 2012 across moneyness levels and time-to-maturities. This is primarily because the underlying VXX index declined over time due to the negative roll yield of VIX futures. Lastly, according to the trade volume and open interest, VIX options are more liquid than VXX options.

We construct the RV from the 1-min high-frequency VIX data, following the procedure in Christoffersen et al. (2014), Feunou and Okou (2019), and Z. Pan et al. (2021). Specifically, on each day, the RV is computed as the sum of 5-min squared returns; there are five daily RVs computed in total: the first RV is computed starting from the first price in the 1-min VIX price grid, and the second RV is computed starting from the second price in the 1-min VIX price grid, and so on, and the RV estimate for the day is the average of the five RVs. Lastly, following Hansen and Lunde (2005), the daily RV series is scaled to match the unconditional variance of the VIX returns over the sample period by using the daily VIX returns. Table 3 reports the summary statistics for the daily RV of the VIX index for both in-sample and out-of-sample. The Jarque–Bera test statistic suggests that the RV is not normally distributed; the RV has a positive skewness and is highly concentrated around its mean, as suggested by the kurtosis. Figure 3 exhibits the evolution of the daily realized volatility—the squared root of daily RV—of the VIX index.

4.2 | Joint Estimation

Models are estimated by maximizing the following joint log-likelihood function over the sample period by using data on both Wednesday VIX options and Wednesday VXX options:

$$\ln \mathcal{L}^{\text{all}} = \text{Scaling constant} \times \left(\frac{\ln \mathcal{L}^{\text{F}}}{N^{\text{F}}} + \frac{\ln \mathcal{L}^{\text{vix}}}{N^{\text{vix}}} + \frac{\ln \mathcal{L}^{\text{vxx}}}{N^{\text{vxx}}} \right) \quad (27)$$

with

$$\ln \mathcal{L}^{\text{F}} = -\frac{N^{\text{F}}}{2} \ln\left(2\pi\sigma_{\text{F}}^2\right) - \frac{1}{2\sigma_{\text{F}}^2} \sum_{n=1}^{N^{\text{F}}} \left(\frac{\widetilde{F}_n^{\text{vix}} - F_n^{\text{vix}}}{F_n^{\text{vix}}} \right)^2,$$

$$\ln \mathcal{L}^{\text{vix}} = -\frac{N^{\text{vix}}}{2} \ln\left(2\pi\sigma_{\text{vix}}^2\right) - \frac{1}{2\sigma_{\text{vix}}^2} \sum_{i=1}^{N^{\text{vix}}} \left(\frac{\widetilde{O}_i^{\text{vix}} - O_i^{\text{vix}}}{O_i^{\text{vix}}} \right)^2,$$

$$\ln \mathcal{L}^{\text{vxx}} = -\frac{N^{\text{vxx}}}{2} \ln\left(2\pi\sigma_{\text{vxx}}^2\right) - \frac{1}{2\sigma_{\text{vxx}}^2} \sum_{j=1}^{N^{\text{vxx}}} \left(\frac{\widetilde{O}_j^{\text{vxx}} - O_j^{\text{vxx}}}{O_j^{\text{vxx}}} \right)^2,$$

where N^{F}, N^{vix}, and N^{vxx} are the total numbers of VIX futures and options and VXX options in the sample, respectively. $\widetilde{F}^{\text{vix}}$ and F^{vix} are model-implied and market-quoted VIX futures prices, respectively; $\widetilde{O}^{\text{vix}}$ and O^{vix} are model-implied and market-quoted VIX option prices, respectively; $\widetilde{O}^{\text{vxx}}$ and O^{vxx} are model-implied and market-quoted VXX option prices,

TABLE 7 | Out-of-sample pricing performance: Overall (estimate-and-forget method).

	MAE			RMSE			MAPE		
	HAR-GARCH	HAR-RV-GARCH	Reduction (%)	HAR-GARCH	HAR-RV-GARCH	Reduction (%)	HAR-GARCH	HAR-RV-GARCH	Reduction (%)
VIX futures	3.4431	**2.9749**	13.60	3.8765	**3.3726**	13.00	0.1939	**0.1675**	13.62
VIX options	0.8380	**0.7323**	12.61	1.2469	**1.0141**	18.67	0.7280	**0.7082**	2.72
VXX options	0.2257	**0.1577**	30.13	0.2994	**0.2026**	32.33	0.3548	**0.2338**	34.10

t statistics		
VIX futures	**VIX options**	**VXX options**
16.6687	**8.1521**	**3.1854**

Note: This table reports the overall out-of-sample pricing performance for the sample period from January 2 to October 10, 2013. Pricing performance is measured by the error metrics, including the Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE), and the Mean Absolute Percentage Error (MAPE). Pricing errors are computed by using parameter estimates in Table 4. The *t* statistics are from the pairwise *t* test of Huang and Wu (2004) on the sample differences in the daily mean squared error. A positive *t* statistic that is larger than 1.645 indicates that the HAR-RV-GARCH model outperforms the HAR-GARCH model at a 5% significance level. Numbers in bold indicate the smallest error statistics or significance at a 5% significance level for the pairwise *t* test.

Abbreviations: GARCH, generalized autoregressive conditionally heteroskedastic; HAR, heterogeneous autoregressive; RV, realized variance; VIX, volatility index.

respectively. The derivatives pricing error is defined as a percentage error of the derivatives price and is assumed to be normally distributed with zero mean and variances of $\sigma^2_{\mathrm{F}}, \sigma^2_{\mathrm{vix}}$, and σ^2_{vxx}, which are estimated with a sample variance of the pricing errors. The percentage error has the advantage of avoiding giving a high weight to NTM options; for example, Ornthanalai (2014), Cao et al. (2020), and Zhou et al. (2024) adopt a similar pricing error function in the model estimation. The weighted log-likelihood function prevents the model parameter estimates from being dominated by a particular type of data; the choice of the scaling constant does not affect the estimation, and we choose to use $(N^{\mathrm{F}} + N^{\mathrm{vix}} + N^{\mathrm{vxx}})/3$ by following Ornthanalai (2014). The starting values for model estimation using the log-likelihood function are taken from those reported by Jiang et al. (2022).

4.3 | Pricing Performance Measures

To quantify the pricing performance, first, commonly used error metrics including the Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE), and the Mean Absolute Percentage Error (MAPE) are used and are computed as follows:

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|\widetilde{O}_i - O_i\right|, \tag{28}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\widetilde{O}_i - O_i)^2}, \tag{29}$$

TABLE 8 | Out-of-sample pricing performance by moneyness and time-to-maturity (estimate-and-forget method).

	MAE		RMSE		MAPE	
	HAR-GARCH	HAR-RV-GARCH	HAR-GARCH	HAR-RV-GARCH	HAR-GARCH	HAR-RV-GARCH
VIX futures						
$\tau \leq 30$	0.7747	**0.6501**	0.8832	**0.7596**	0.0509	**0.0434**
$30 < \tau \leq 90$	2.2987	**1.9194**	2.4581	**2.0522**	0.1373	**0.1148**
$\tau > 90$	4.8488	**4.2382**	4.9427	**4.3284**	0.2662	**0.2326**
VIX options						
$m \leq -0.2$	**0.1397**	0.2233	**0.1886**	0.2812	**0.3100**	0.4552
$-0.2 < m \leq -0.06$	**0.4219**	0.5527	**0.5326**	0.6692	**0.3667**	0.4463
$-0.06 < m \leq 0.06$	1.5719	**1.3380**	1.9920	**1.6239**	0.7206	**0.6136**
$0.06 < m \leq 0.2$	1.6983	**1.2629**	2.0133	**1.5314**	1.0116	**0.7555**
$0.2 < m \leq 0.4$	0.9641	**0.7699**	1.2551	**0.9906**	0.8749	**0.7226**
$0.4 < m \leq 0.6$	0.4941	**0.4775**	0.6780	**0.6171**	0.7380	**0.7410**
$m > 0.6$	**0.2210**	0.3275	**0.3109**	0.4041	**0.6555**	1.0438
$\tau \leq 30$	0.2225	**0.1555**	0.3435	**0.2382**	0.4804	**0.3287**
$30 < \tau \leq 90$	0.5075	**0.4348**	0.7643	**0.5810**	0.5045	**0.4947**
$\tau > 90$	1.3049	**1.1576**	1.6868	**1.3915**	1.0071	**1.0060**
VXX options						
$m \leq -0.2$	0.4248	**0.1452**	0.4469	**0.1681**	0.9684	**0.3143**
$-0.2 < m \leq -0.06$	0.4063	**0.1693**	0.4534	**0.1979**	0.7233	**0.2895**
$-0.06 < m \leq 0.06$	0.3087	**0.1679**	0.3770	**0.1987**	0.2510	**0.1565**
$0.06 < m \leq 0.2$	0.1937	**0.1535**	0.2686	**0.1896**	0.2023	**0.1966**
$0.2 < m \leq 0.4$	**0.1475**	0.1588	**0.2057**	0.2077	**0.2048**	0.2097
$0.4 < m \leq 0.6$	**0.1346**	0.1595	**0.1903**	0.2209	**0.2338**	0.2438
$m > 0.6$	**0.1288**	0.1486	**0.1609**	0.2175	0.3307	**0.2903**
$\tau \leq 30$	0.1868	**0.1059**	0.2733	**0.1338**	0.4877	**0.2946**
$30 < \tau \leq 90$	0.1990	**0.1263**	0.2699	**0.1525**	0.3916	**0.2581**
$\tau > 90$	0.2673	**0.2101**	0.3322	**0.2591**	0.2524	**0.1814**

Note: This table reports the out-of-sample pricing performance by option moneyness and time-to-maturity for the sample period from January 2 to October 10, 2013. Error metrics, including the Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE), and the Mean Absolute Percentage Error (MAPE), are used. Pricing errors are computed by using parameter estimates in Table 4. Option moneyness m is defined as the logarithmic of the ratio between the strike price and the underlying forward price. τ denotes the time-to-maturity (days). Numbers in bold indicate the smallest error statistics.

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\frac{|\widetilde{O}_i - O_i|}{O_i}. \tag{30}$$

Second, to measure the statistical significance of model j over model i in terms of pricing, following Huang and Wu (2004), we perform the pairwise t test on the sample differences in the daily mean squared error, and the test statistic is computed as follows:

$$t = \frac{\bar{\text{MSE}}_i - \bar{\text{MSE}}_j}{\text{stdev}(\bar{\text{MSE}}_i - \bar{\text{MSE}}_j)}, \tag{31}$$

wherein the nominator is the difference in the sample means of daily mean squared errors; the denominator is the standard error of the difference, and it is adjusted for serial correlation following Newey and West (1987), where the optimal lag number is chosen according to Andrews (1991) and an AR(1) specification. When $t > 1.645$, model j significantly outperforms model i at the 5% significance level.

5 | Results

5.1 | Parameter Estimates

The first panel in Table 4 reports the parameter estimates and standard errors from the maximum log-likelihood estimation over the in-sample period from January 3 to December 31, 2012. Most parameter estimates are statistically significant at a 5% significance level. For both models, the coefficient of the 1-day lagged VIX observation (β_d) is much larger compared with the coefficients of 1-week and 1-month lagged VIX values (β_w and β_m), implying that the 1-day lagged VIX index contributes more to the joint pricing of VIX futures and options and VXX options. Compared with estimates of β_d and β_m, β_w estimates are considerably smaller, this is consistent with estimates in the prior studies, when VIX derivatives data are used to estimate the parameters of the HAR-GARCH models that are directly used to model the VIX index.[7] The persistence estimate for the HAR-RV-GARCH (0.8990) is slightly lower than that for the HAR-GARCH model (0.9035), but both persistence estimates are large, indicating that the conditional variance of the daily log-VIX is highly persistent under $\mathbb{Q}$.

The second and third panels in Table 4 report the average log-likelihood per contract at the optimum and the information criteria for VIX futures and options and VXX options pricing errors. Akaike information criterion (AIC) and Bayesian information criterion (BIC) are used and are calculated as $\text{AIC} = (2k - 2\ln\mathcal{L})/N$, $\text{BIC} = (k\ln N - 2\ln\mathcal{L})/N$, where k is the number of model parameters, N is the total number of derivatives contracts, $\ln\mathcal{L}$ is the log-likelihood function value. The average log-likelihood function value from the HAR-RV-GARCH model is larger than that from the HAR-GARCH model for both the VIX derivatives and VXX options; by conducting a likelihood ratio test, the pricing performance of the HAR-RV-GARCH model is statistically significantly better than that of the HAR-GARCH model for the joint pricing in-sample. The information criteria also support the findings.

TABLE 9 | Out-of-sample pricing performance: Overall (rolling window method).

	MAE			RMSE			MAPE		
	HAR-GARCH	HAR-RV-GARCH	Reduction (%)	HAR-GARCH	HAR-RV-GARCH	Reduction (%)	HAR-GARCH	HAR-RV-GARCH	Reduction (%)
VIX futures	0.8674	**0.7956**	8.28	1.2834	**1.1820**	7.90	0.0496	**0.0455**	8.27
VIX options	0.3089	**0.2214**	28.33	0.4547	**0.3398**	25.27	0.4100	**0.2469**	39.78
VXX options	0.2136	**0.1147**	46.30	0.2734	**0.1620**	40.75	0.3647	**0.1689**	53.69

t statistics		
VIX futures	VIX options	VXX options
2.8178	**2.9664**	**3.7136**

Note: This table reports the overall out-of-sample pricing performance for the sample period from January 2 to October 10, 2013. Pricing performance is measured by the error metrics, including the Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE), and the Mean Absolute Percentage Error (MAPE). Model parameters are updated every month with a 6-month estimation window; the estimated model parameters are then used to price VIX derivatives and VXX options in the month that succeeds the 6-month estimation window. The t statistics are from the pairwise t test of Huang and Wu (2004) on the sample differences in the daily mean squared error. A positive t statistic that is larger than 1.645 indicates that the HAR-RV-GARCH model outperforms the HAR-GARCH model at a 5% significance level. Numbers in bold indicate the smallest error statistics or significance at a 5% significance level for the pairwise t test.
Abbreviations: GARCH, generalized autoregressive conditionally heteroskedastic; HAR, heterogeneous autoregressive; RV, realized variance; VIX, volatility index.

5.2 | In-Sample Pricing Errors Analysis

Table 5 reports the overall in-sample pricing performance measured by error metrics and the pairwise *t* test. In terms of the pricing error, the model that incorporates the RV of the VIX (HAR-RV-GARCH) produces smaller error statistics in all categories than the model that does not rely on VIX RV (HAR-GARCH), except for VIX futures, according to MAE. The performance of pricing VIX futures in-sample is close. HAR-RV-GARCH only reduces RMSE by 0.82% and MAPE by 1.35%; however, HAR-RV-GARCH reduces option pricing errors by a much larger extent: for example, HAR-RV-GARCH, as compared with the HAR-GARCH, reduces MAPE by 30.58% and 35.03% for pricing VIX and VXX options, respectively.

To supplement the error statistics, the pairwise *t* test of Huang and Wu (2004) assesses the statistical significance of the outperformance of one model over the other. A positive test statistic that is larger than 1.645 reported in the table indicates that the HAR-RV-GARCH model outperforms the HAR-GARCH model significantly at a 5% significance level. The *t* statistic for VIX futures pricing is positive, indicating that the HAR-RV-GARCH outperforms the HAR-GARCH, though the outperformance is insignificant, whereas for VIX and VXX options pricing, the HAR-RV-GARCH significantly outperforms the HAR-GARCH, with *t* statistics standing at 3.3184 and 5.2285, respectively.

Table 6 reports the in-sample pricing performance by moneyness and time-to-maturity. In terms of VIX futures pricing,

TABLE 10 | Out-of-sample pricing performance by moneyness and time-to-maturity (rolling window method).

	MAE		RMSE		MAPE	
	HAR-GARCH	HAR-RV-GARCH	HAR-GARCH	HAR-RV-GARCH	HAR-GARCH	HAR-RV-GARCH
VIX futures						
$\tau \leq 30$	0.4306	**0.3529**	0.5365	**0.4352**	0.0279	**0.0228**
$30 < \tau \leq 90$	0.6491	**0.5844**	0.8878	**0.8084**	0.0392	**0.0354**
$\tau > 90$	1.1137	**1.0387**	1.5907	**1.4728**	0.0616	**0.0576**
VIX options						
$m \leq -0.2$	0.2378	**0.1188**	0.2702	**0.1568**	0.7166	**0.2873**
$-0.2 < m \leq -0.06$	**0.2299**	0.2667	**0.2834**	0.3614	0.3300	**0.2568**
$-0.06 < m \leq 0.06$	0.4472	**0.4130**	0.7002	**0.5822**	0.2401	**0.2067**
$0.06 < m \leq 0.2$	0.4814	**0.3375**	0.7178	**0.4805**	0.3393	**0.2237**
$0.2 < m \leq 0.4$	0.3230	**0.2090**	0.4226	**0.2828**	0.3687	**0.2321**
$0.4 < m \leq 0.6$	0.2424	**0.1337**	0.2749	**0.1726**	0.4501	**0.2519**
$m > 0.6$	0.1684	**0.0896**	0.1861	**0.1191**	0.5748	**0.3039**
$\tau \leq 30$	0.2450	**0.1417**	0.3224	**0.1803**	0.5896	**0.3187**
$30 < \tau \leq 90$	0.2865	**0.1854**	0.3692	**0.2517**	0.4369	**0.2427**
$\tau > 90$	0.3455	**0.2749**	0.5478	**0.4313**	0.3439	**0.2356**
VXX options						
$m \leq -0.2$	0.2451	**0.0944**	0.2931	**0.1240**	0.6003	**0.1778**
$-0.2 < m \leq -0.06$	0.2440	**0.1262**	0.3121	**0.1562**	0.4894	**0.2101**
$-0.06 < m \leq 0.06$	0.2127	**0.1325**	0.3053	**0.1676**	0.2109	**0.1176**
$0.06 < m \leq 0.2$	0.1623	**0.1143**	0.2304	**0.1490**	0.2096	**0.1493**
$0.2 < m \leq 0.4$	0.1988	**0.1068**	0.2468	**0.1476**	0.2892	**0.1558**
$0.4 < m \leq 0.6$	0.2363	**0.1080**	0.2854	**0.1659**	0.3968	**0.1623**
$m > 0.6$	0.2338	**0.1201**	0.2685	**0.2073**	0.5499	**0.2352**
$\tau \leq 30$	0.1617	**0.0990**	0.2459	**0.1300**	0.4526	**0.2538**
$30 < \tau \leq 90$	0.2046	**0.0902**	0.2550	**0.1195**	0.4392	**0.1662**
$\tau > 90$	0.2496	**0.1415**	0.2999	**0.2006**	0.2605	**0.1227**

Note: This table reports the out-of-sample pricing performance by option moneyness and time-to-maturity for the sample period from January 2 to October 10, 2013. Error metrics, including the Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE), and the Mean Absolute Percentage Error (MAPE), are used. Model parameters are updated every month with a 6-month estimation window; the estimated model parameters are then used to price VIX derivatives and VXX options in the month that succeeds the 6-month estimation window. Option moneyness m is defined as the logarithmic of the ratio between the strike price and the underlying forward price. τ denotes the time-to-maturity (days). Numbers in bold indicate the smallest error statistics.
Abbreviations: GARCH, generalized autoregressive conditionally heteroskedastic; HAR, heterogeneous autoregressive; RV, realized variance; VIX, volatility index.

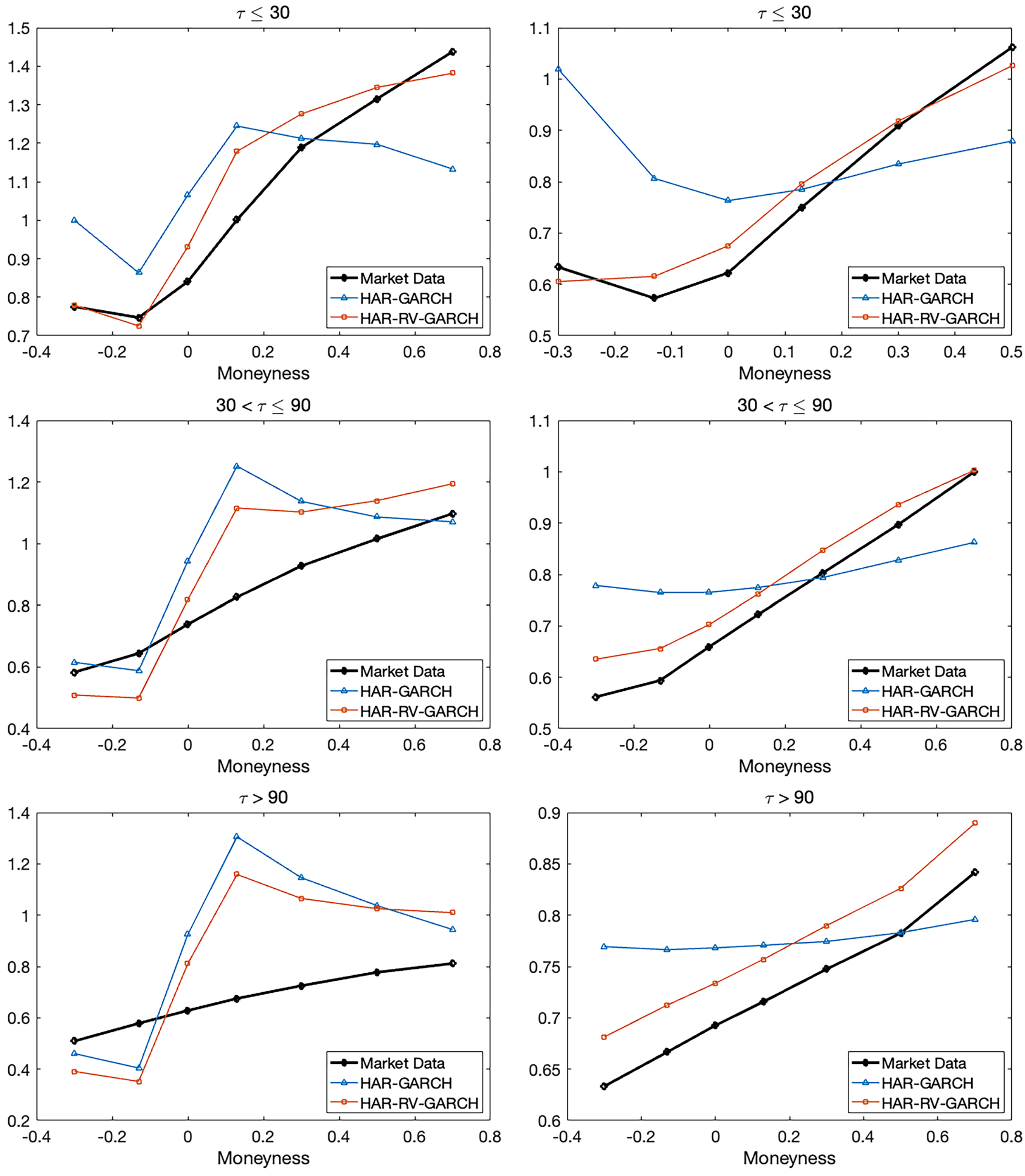

FIGURE 5 | Out-of-sample fit to the implied volatility (estimate-and-forget method). The figure exhibits the average out-of-sample fit to the VIX and VXX implied volatilities across moneyness and time-to-maturity for the sample period from January 2 to October 10, 2013. Figures in the left column show the average out-of-sample fit to the VIX implied volatility, and figures in the right column show the average out-of-sample fit to the VXX implied volatility. Option moneyness is defined as the logarithm of the ratio between the strike price of the option and the underlying asset forward price. GARCH, generalized autoregressive conditionally heteroskedastic; HAR, heterogeneous autoregressive; RV, realized variance; VIX, volatility index. [Color figure can be viewed at wileyonlinelibrary.com]

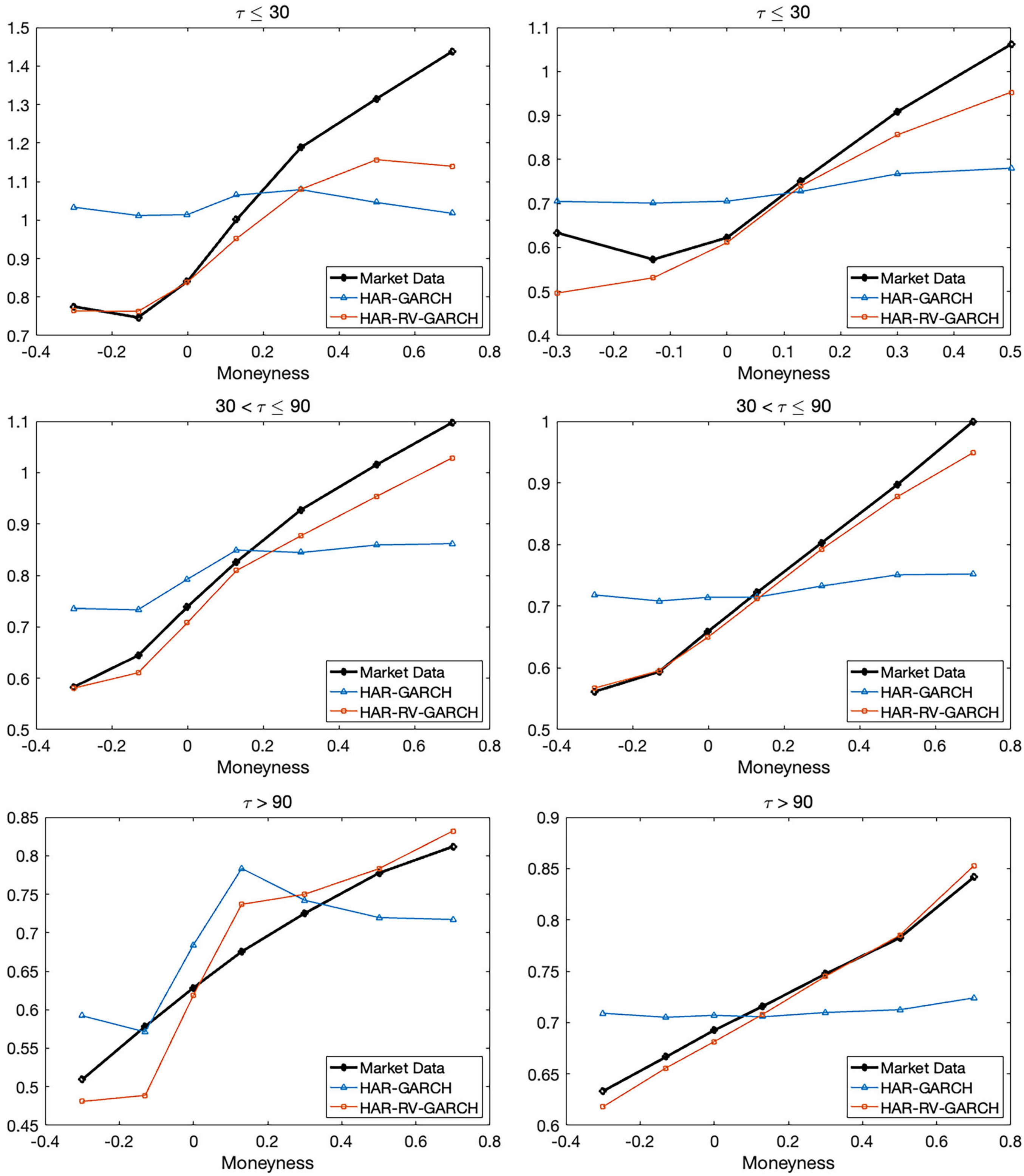

FIGURE 6 | Out-of-sample fit to the implied volatility (rolling window method). The figure exhibits the average out-of-sample fit to the VIX and VXX implied volatilities across moneyness and time-to-maturity for the sample period from January 2 to October 10, 2013. Figures in the left column show the average out-of-sample fit to the VIX implied volatility, and figures in the right column show the average out-of-sample fit to the VXX implied volatility. Option moneyness is defined as the logarithm of the ratio between the strike price of the option and the underlying asset forward price. GARCH, generalized autoregressive conditionally heteroskedastic; HAR, heterogeneous autoregressive; RV, realized variance; VIX, volatility index. [Color figure can be viewed at wileyonlinelibrary.com]

while the HAR-GARCH model performs better for short- and medium-term contracts, HAR-RV-GARCH is better at pricing long-term contracts. Moving to the VIX option pricing, first, regarding the pricing performance by option moneyness, the HAR-RV-GARCH outperforms the HAR-GARCH in pricing deep-OTM puts and OTM calls, while it is inferior to HAR-GARCH for NTM options; second, in terms of the pricing performance by time-to-maturity, the HAR-RV-GARCH outperforms the HAR-GARCH over all maturities. Turning to the VXX option pricing, similar findings are observed: First, the

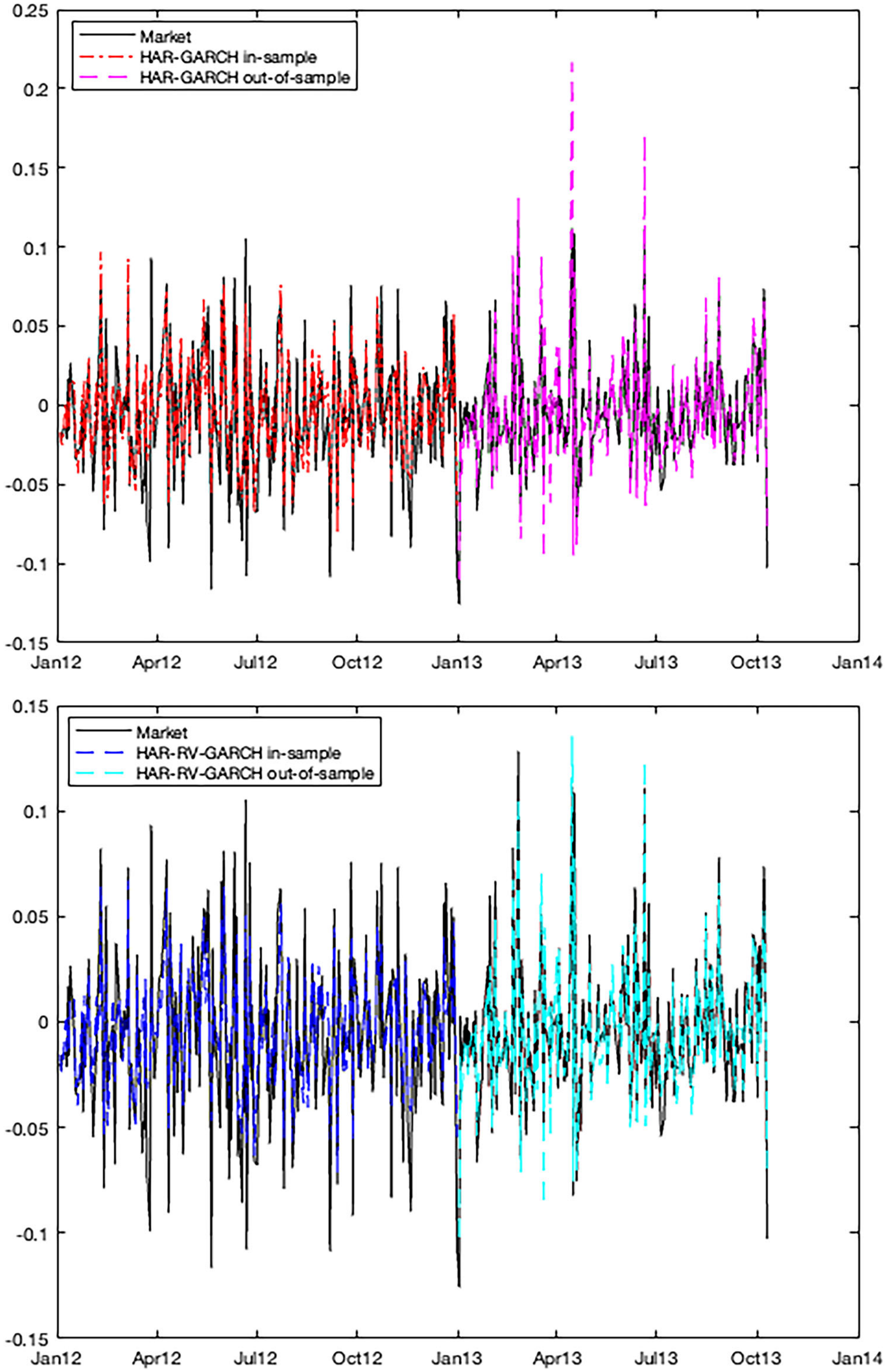

FIGURE 7 | Model fit to VXX return dynamics. The figure exhibits the model fit to the VXX return dynamics for the period from January 4, 2012 to October 10, 2013. VXX return estimates are obtained by using parameter estimates in Table 4, and using (16) for the HAR-RV-GARCH model and (20) for the HAR-GARCH model. GARCH, generalized autoregressive conditionally heteroskedastic; HAR, heterogeneous autoregressive; RV, realized variance. [Color figure can be viewed at wileyonlinelibrary.com]

HAR-RV-GARCH model is superior over the HAR-GARCH model in pricing OTM VXX calls and puts, while it underperforms the HAR-GARCH in pricing NTM VXX options; second, HAR-RV-GARCH outperforms the HAR-GARCH in pricing VXX options over all maturity groups.

Figure 4 exhibits the average in-sample fit to the VIX and VXX option implied volatilities across different moneyness levels. Figures in the left column show the average fit to the VIX option implied volatility skew, while figures in the right column show the average fit to the implied volatility skew of VXX options. The figures confirm our findings reported in Table 6.

5.3 | Out-of-Sample Performance

We carry out out-of-sample pricing analysis to alleviate the concerns that the in-sample pricing performance is driven by overfitting the data rather than by utilizing the information contained in the realized volatility from high-frequency VIX index data.

The first method we adopt is the estimate-and-forget method, where parameter estimates in Table 4 obtained from the in-sample data (2012) are used to price VIX and VXX derivatives in the out-of-sample (2013).

The second method employed is the rolling window method by following Yin et al. (2021), Tong et al. (2022), and T. Wang et al. (2022): model parameters are updated every month with a 6-month estimation window; the estimated model parameters are then used to price VIX derivatives and VXX options in the month (in the out-of-sample) succeeds the 6-month estimation window.[8] For example, model parameters are estimated by using data from July to December 2012, to price derivatives in January 2013 (the first month in the out-of-sample). Compared with the estimate-and-forget method, the rolling window method is more practically relevant, as in practice, traders are more likely to re-estimate and update the model parameters very often and periodically, rather than to estimate the model once and leave it unchanged for a substantial long period since market conditions change.

Table 7 reports the overall out-of-sample pricing performance by using the estimate-and-forget method. Consistent with the findings for the in-sample, HAR-RV-GARCH outperforms HAR-GARCH to a substantial degree according to all the error metrics. The outperformance of the HAR-RV-GARCH over the HAR-GARCH is significant at a 5% significance level, based on the t statistics. Table 8 reports the out-of-sample performance by moneyness and time-to-maturity, using the estimate-and-forget method. The HAR-RV-GARCH produces smaller pricing errors across all maturities. In pricing VIX options, the HAR-RV-GARCH performs better in pricing NTM options and OTM calls; in pricing VXX options, the HAR-RV-GARCH model outperforms the HAR-GARCH model for OTM puts and NTM options.

Using the rolling window method, Table 9 reports the overall out-of-sample performance, and Table 10 reports the out-of-sample performance by moneyness and time-to-maturity. Similar to the findings using the estimate-and-forget method, overall, the HAR-RV-GARCH again significantly outperforms the HAR-GARCH based on the t statistics; the superiority of the HAR-RV-GARCH model over the HAR-GARCH model is evidenced by the smaller pricing errors across all time-to-maturities and almost all the moneyness levels.

The overall improvement of HAR-RV-GARCH over HAR-GARCH is because realized volatility of the high-frequency VIX index allows the HAR-RV-GARCH model to adapt more rapidly to the changes in the volatility of underlying assets, especially in the tail region of the moneyness, and to exploit the information in the high-frequency data more thoroughly than the filtering procedure applied by the HAR-GARCH model across all option maturities (Corsi et al. 2013). While RV from the high-frequency VIX index is an accurate measure of the volatility of VIX and VIX futures, RV from the high-frequency VIX is a much noisier measure of daily variations in the VXX. In fact, Bergomi (2015) argued that the information on vanillas (VIX) could not be used to fully recover the information contained in their exotic path-dependent products (VXX), and Bašta and Molnár (2019) found that information from the VIX index has limited ability to improve forecasts of VXX. We can reasonably conjecture that when the forecasting horizon is long (in the estimate-and-forget exercise, we are forecasting option prices

TABLE 11 | VXX return dynamics: Summary statistics and model fit.

	Summary statistics					
VXX return	**Mean**	**Std. deviation**	**Skewness**	**Kurtosis**	**Minimum**	**Maximum**
In-sample (2012)	− 0.0058	0.0409	− 0.0074	3.0794	− 0.1166	0.1054
Out-of-sample (2013)	− 0.0039	0.0349	0.6550	5.5203	− 0.1258	0.1284
	In-sample fit			**Out-of-sample fit**		
	MAE	**RMSE**	**MAPE**	**MAE**	**RMSE**	**MAPE**
HAR-GARCH	0.0169	**0.0214**	1.5996	0.0117	0.0163	3.2068
HAR-RV-GARCH	**0.0160**	0.0226	**1.4349**	**0.0104**	**0.0132**	**2.7899**

Note: This table reports the summary statistics of the VXX return and the performance of the model fit to the VXX return dynamics for the in-sample period from January 4 to December 31, 2012 and for the out-of-sample period from January 2 to October 10, 2013. VXX return estimates are obtained by using parameter estimates in Table 4, and using (16) for the HAR-RV-GARCH model and (20) for the HAR-GARCH model. Numbers in bold indicate the smallest error statistics. "Std. deviation" stands for the standard deviation.
Abbreviations: GARCH, generalized autoregressive conditionally heteroskedastic; HAR, heterogeneous autoregressive; MAE, Mean Absolute Error; MAPE, Mean Absolute Percentage Error; RMSE, Root Mean Squared Error; RV, realized variance.

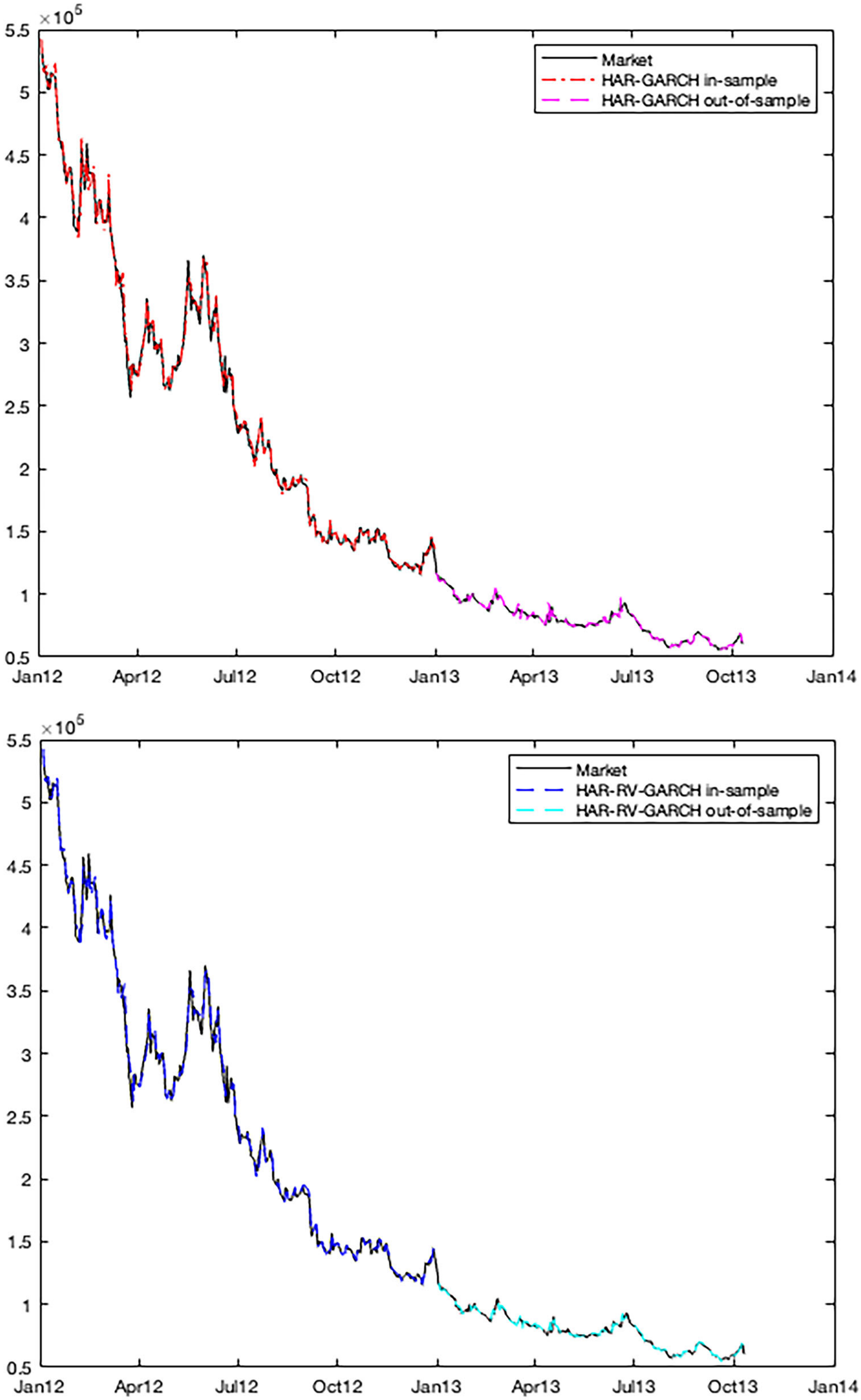

FIGURE 8 | Model fit to VXX price dynamics (1-day ahead). The figure exhibits the model fit to the VXX price dynamics for the in-sample period from January 4 to December 31, 2012, and for the out-of-sample period from January 2 to October 10, 2013. The VXX price is adjusted for reverse splits. The VXX price estimates are 1-day ahead estimates obtained by using (32), where VXX return estimates are obtained by using parameter estimates in Table 4, and using (16) for the HAR-RV-GARCH model and (20) for the HAR-GARCH model. GARCH, generalized autoregressive conditionally heteroskedastic; HAR, heterogeneous autoregressive; RV, realized variance. [Color figure can be viewed at wileyonlinelibrary.com]

9 months into the future), RV may be still informative about extreme movements in the VIX and VIX futures (which is usually associated with increases in the VIX), that is why HAR-RV-GARCH still outperforms HAR-GARCH in pricing OTM VIX call options in the estimate-and-forget out-of-sample exercise (as shown in Table 8), which is consistent with the in-sample analysis (as show in Table 6); whereas RV from VIX may be less informative about the extreme VXX movements far into the future due to the noises, the HAR-RV-GARCH model transmits this noise onto the VXX option price forecasts, yielding less accurate estimates of VXX options, creating distortions in the relative model performance and performance patterns that are inconsistent with the in-sample analysis; when the forecasting horizon is short (in the rolling window exercise, we are forecasting option prices only 1 month into the future), the impact of the aforementioned noises on the RV's ability to forecast future VIX and VXX movements is much smaller, resulting in HAR-RV-GARCH model outperforming the HAR-GARCH across all moneyness of VIX and VXX options (as shown in Table 10).

Figures 5 and 6 exhibit the average out-of-sample fit to the implied volatilities of VIX and VXX options by moneyness and time-to-maturity, using the estimate-and-forget and rolling window method, respectively. The figures in the left column show the fit to the VIX option implied volatility, and figures in the right column show the fit to the VXX option implied volatility. These figures confirm the prior findings.

5.4 | Model Fit to VXX Dynamics

We now turn to the model fit to the VXX dynamics. First, we compute the fitted VXX returns for both the in-sample and out-of-sample periods by using the model parameter estimates reported in Table 4. We use (16) to compute the VXX return estimates under the HAR-RV-GARCH model and use (20) to compute VXX return estimates under the HAR-GARCH model. Figure 7 shows the model fit to the VXX return dynamics from January 2012 to October 2013. In general, both models' VXX return estimates track the market VXX returns reasonably close, both in-sample and out-of-sample; however, it is difficult to assess the performance only from the figure. As a result, we subsequently compute the error statistics to quantify the model fit to the VXX return, and the results are reported in Table 11. It shows that the HAR-RV-GARCH model performs better than the HAR-GARCH model in fitting the VXX return dynamics both in-sample and out-of-sample.

Second, we use the VXX return estimates from the first step to compute the VXX price estimates to assess the model fit to the VXX price dynamics both in-sample and out-of-sample. Given the return estimate for day $t + 1$ ($\widetilde{R_{t+1}^{\text{VXX}}}$) and the market VXX price observed on day t (the previous day), the estimate for VXX price on day $t + 1$ is computed as

$$\widetilde{\text{VXX}_{t+1}} = \text{VXX}_t \cdot e^{\widetilde{R_{t+1}^{\text{VXX}}}}. \tag{32}$$

And we plot the VXX price 1-day ahead estimates $\{\widetilde{\text{VXX}_{t+1}}\}$ under both models in Figure 8. It is clear that both HAR-RV-GARCH and HAR-GARCH models fit the VXX price dynamics extremely well, both in-sample and out-of-sample. The top and middle panels in Table 12 report the summary statistics of and 1-day ahead model fit to the VXX price dynamics. The MAPE stands at 1.61% and 1.7% in-sample and 1.17% and 1.04% out-of-sample for HAR-GARCH and HAR-RV-GARCH, respectively, indicating extremely good performance of discrete-time VXX models in fitting VXX price dynamics. The performance of the two models is very close, with HAR-RV-GARCH slightly better.

Third, instead of computing one-step ahead estimates, we compute multistep ahead VXX price estimates for the out-of-

TABLE 12 | VXX price dynamics: Summary statistics and model fit.

	Summary statistics					
VXX price	**Mean**	**Std. deviation**	**Skewness**	**Kurtosis**	**Minimum**	**Maximum**
In-sample (2012)	259,320	112,820	0.5620	2.2832	115,310	540,840
Out-of-sample (2013)	78,939	14,606	0.3356	2.4431	55,624	115,220

	In-sample fit			**Out-of-sample fit**		
	MAE	**RMSE**	**MAPE**	**MAE**	**RMSE**	**MAPE**
One-day ahead						
HAR-GARCH	4352.8	6179.6	0.0170	939.1	1373.8	0.0117
HAR-RV-GARCH	**4114.1**	**5901.2**	**0.0161**	**832.5**	**1086.4**	**0.0104**
Multistep ahead						
HAR-GARCH	—	—	—	7961.8	8504.8	0.1080
HAR-RV-GARCH	—	—	—	**5939.6**	**6393.4**	**0.0797**

Note: This table reports the summary statistics of the VXX price and the performance of the model fit to the VXX price dynamics for the in-sample period from January 4 to December 31, 2012, and for the out-of-sample period from January 2 to October 10, 2013. The VXX price is adjusted for reverse splits. The 1-day ahead VXX price estimates are obtained by using (32), and the multistep ahead VXX price estimates are obtained by using (33) and the market-observed VXX price on December 31, 2012 (reverse-split adjusted value), where VXX return estimates are obtained by using parameter estimates in Table 4, and using (16) for the HAR-RV-GARCH model and (20) for the HAR-GARCH model. Numbers in bold indicate the smallest error statistics. "Std. deviation" stands for the standard deviation.
Abbreviations: GARCH, generalized autoregressive conditionally heteroskedastic; HAR, heterogeneous autoregressive; MAE, Mean Absolute Error; MAPE, Mean Absolute Percentage Error; RMSE, Root Mean Squared Error; RV, realized variance.

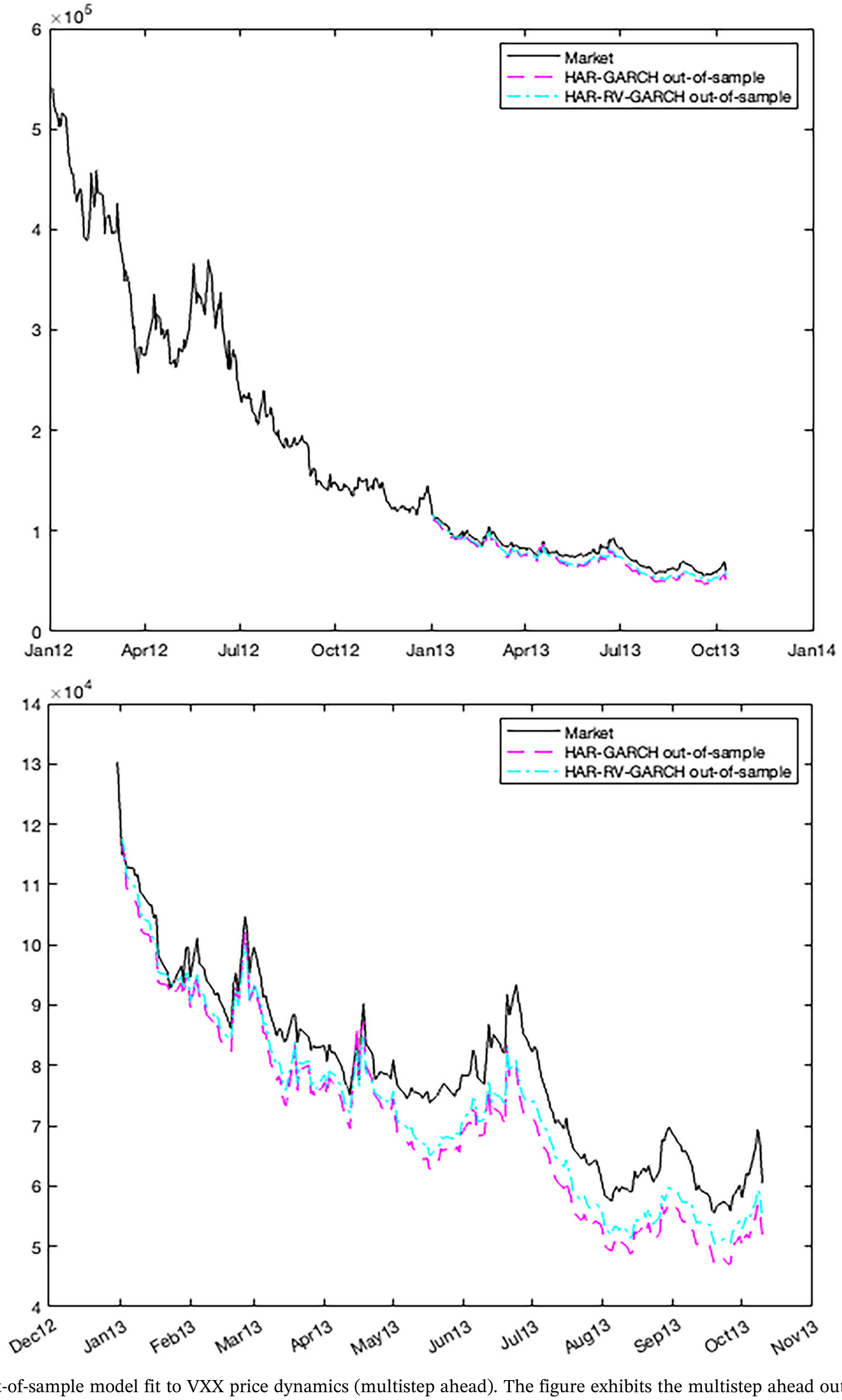

FIGURE 9 | Out-of-sample model fit to VXX price dynamics (multistep ahead). The figure exhibits the multistep ahead out-of-sample model fit to the VXX price dynamics for the out-of-sample period from January 2 to October 10, 2013. The top figure shows the out-of-sample model fit; the bottom figure shows a close-up of the out-of-sample model fit that is displayed in the top figure. The VXX price is adjusted for reverse splits. The VXX price estimates in the out-of-sample are obtained by using (33) and the market-observed VXX price on December 31, 2012 (reverse-split adjusted value), where VXX return estimates are obtained by using parameter estimates in Table 4, and using (16) for the HAR-RV-GARCH model and (20) for the HAR-GARCH model. GARCH, generalized autoregressive conditionally heteroskedastic; HAR, heterogeneous autoregressive; RV, realized variance. [Color figure can be viewed at wileyonlinelibrary.com]

sample from January 2 to October 10, 2013. We only use market-observed VXX price on December 31, 2012 (the last day in the in-sample period), denoted as VXX_0, and we know that $\text{VXX}_0 = 130,293.76$ (reverse-split adjusted value), and compute the VXX price estimates in the out-of-sample according to the equation below:

$$\widetilde{\text{VXX}}_n = \text{VXX}_0 \prod_{i=1}^{n} e^{\widetilde{R_n^{\text{VXX}}}}, \quad (33)$$

where $\widetilde{\text{VXX}}_n$ is the VXX price estimate on the nth day in the out-of-sample, $\widetilde{R_n^{\text{VXX}}}$ is the VXX return estimate on the nth day in the out-of-sample obtained from the first step. The top figure of Figure 9 shows the out-of-sample multistep ahead model fit to the VXX price dynamics; the bottom figure of Figure 9 provides a close-up of the out-of-sample model fit that is displayed in the top figure. It shows that multistep ahead price estimates produced by HAR-RV-GARCH are much closer to the market price than those produced by HAR-GARCH. The bottom panel in Table 12 reports the error statistics for the out-of-sample multistep ahead model fit. Both models fit the price dynamics quite well, with MAPE standing at 10.8% and 7.97% for the HAR-GARCH and HAR-RV-GARCH, respectively. The error statistics again indicate a better performance of the HAR-RV-GARCH model over the HAR-GARCH model.

6 | Conclusion

In this paper, we develop a discrete-time joint analytical framework for consistently pricing VIX and VXX options. Compared with the continuous-time VXX models, the discrete-time pricing framework is more flexible as it allows the information contained in the high-frequency VIX index data to be incorporated for the joint pricing, which the continuous-time VXX option pricing models in the literature are incapable of.

An empirical analysis is performed to assess the models' pricing performance and, in particular, to examine the information content of the RV computed from the high-frequency VIX index data for the joint pricing of VIX derivatives and VXX options. We find that the model that utilizes the VIX RV from the high-frequency VIX data significantly outperforms the model that does not rely on the RV measure in joint pricing, implying that the high-frequency VIX data contain valuable information about the VIX and VXX option pricing.

However, we must point out that the joint pricing performance of the discrete-time VXX models is not completely satisfactory, as the pricing errors are still quite substantial. There are two reasons: First, the discrete-time models are driven by one source of randomness, compared with continuous-time models where there are usually several sources of randomness, such as multifactor VXX models in Tan et al. (2021), Cao et al. (2021), and Lu (2023). Second, our models do not assume jumps in the underlying VIX index and the implied VXX dynamics. There is a consensus that jumps play an important role in option valuation; this is evidenced by the findings of Bao et al. (2012), Tan et al. (2021), Cao et al. (2021), and Lu (2023) in the VXX option pricing literature.

The above discussion indicates some directions for future research in the joint pricing of VIX and VXX options. First, the discrete-time VXX option pricing model could be refined by adding jumps in the assumed VIX dynamics. Second, in addition to the realized volatility computed from the high-frequency VIX index data, future work could exploit the information content of the realized jump variations from the high-frequency VIX about the joint pricing of VIX and VXX options. Third, future work could also explore whether separate modeling of the realized upside and downside volatility dynamics of the VIX index improves the discrete-time model's performance in joint pricing.

Conflicts of Interest

The author declares no conflicts of interest.

Data Availability Statement

The data employed in this paper are available from the CBOE website, Bloomberg terminal, and the FirstRate data website. Restrictions apply to the availability of these data, which were used under license for this study.

Endnotes

[1] This is why both series A and B of VXX ETN are now usually referred to as VXX for short, without distinguishing between series A and B.

[2] VXX option trade volume data are from the CBOE website: https://www.cboe.com/us/options/market_statistics/historical_data/.

[3] It should be noted that although the notations for H functions in (18) and (22) are the same, H functions are different under the HAR-RV-GARCH and HAR-GARCH models.

[4] FirstRate Data: https://firstratedata.com/.

[5] We were restrained by limited resources from getting access to the VXX options data for a considerably longer period.

[6] This practice is due to the fact that the VXX does not pay any dividends or coupons, and the market prices VXX options using the Black–Scholes implied volatility surface of the VXX (Grasselli and Wagalath 2020).

[7] See, for example, Jiang et al. (2022, Tables 3a and 3b), Qiao and Jiang (2023, Table 3a and 3b), and Guo et al. (2024, Tables 4 and 5).

[8] We also used an alternative 1-year estimation window for the out-of-sample pricing exercise in the unreported results, and the relative performance of the models does not change.

References

Alexander, C., J. Kapraun, and D. Korovilas. 2015. "Trading and Investing in Volatility Products." *Financial Markets, Institutions and Instruments* 24: 313–347.

Alexander, C., and D. Korovilas. 2013. "Volatility Exchange-Traded Notes: Curse or Cure?" *Journal of Alternative Investments* 16, no. 2: 52–70.

Alexander, C., D. Korovilas, and J. Kapraun. 2016. "Diversification With Volatility Products." *Journal of International Money and Finance* 65: 213–235.

Alitab, D., G. Bormetti, F. Corsi, and A. Majewski. 2020. "A Jump and Smile Ride: Jump and Variance Risk Premia in Option Pricing." *Journal of Financial Econometrics* 18, no. 1: 121–157.

Andrews, D. 1991. "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation." *Econometrica* 59: 817–858.

Augustyniak, M., A. Badescu, J. Begin, and S. Jayaraman. 2025. "A General Option Pricing Framework for Affine Factionally Integrated Models." *Journal of Banking and Finance* 171: 107346.

Bao, Q., S. Li, and D. Gong. 2012. "Pricing VXX Option With Default Risk and Positive Volatility Skew." *European Journal of Operational Research* 223: 246–255.

Bašta, M., and P. Molnár. 2019. "Long-Term Dynamics of the VIX Index and Its Tradable Counterpart VXX." *Journal of Futures Markets* 39: 322–341.

Bergomi, L. 2015. *Stochastic Volatility Modeling*. CRC Press.

Bu, R., R. Hizmeri, M. Izzeldin, and A. Murphy. 2023. "The Contribution of Jump Signs and Activity to Forecasting Stock Price Volatility." *Journal of Empirical Finance* 70: 144–164.

Cao, H., A. Badescu, Z. Cui, and S. Jayaraman. 2020. "Valuation of VIX and Target Volatility Options With Affine GARCH Models." *Journal of Futures Markets* 40: 1880–1917.

Cao, J., X. Ruan, S. Su, and W. Zhang. 2021. "Specification Analysis of VXX Option Pricing Models Under Lévy Processes." *Journal of Futures Markets* 41: 1456–1477.

Carr, P., and L. Wu. 2010. "Stock Options and Credit Default Swaps: A Joint Framework for Valuation and Estimation." *Journal of Financial Econometrics* 8: 409–449.

CBOE VIX Futures. 2023. [data-set] https://www.cboe.com/us/futures/market_statistics/historical_data/.

CBOE VIX Options. 2023. [data-set] https://datashop.cboe.com/.

CBOE VXX Options. 2023. [data-set] https://datashop.cboe.com/.

CBOE VXX Options Trade Volume. 2024. [data-set] https://www.cboe.com/us/options/market_statistics/historical_data/.

Chernov, M., and E. Ghysels. 2000. "A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation." *Journal of Financial Economics* 56: 407–458.

Christoffersen, P., S. Heston, and K. Jacobs. 2009. "The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work so Well." *Management Science* 55: 1914–1932.

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* 3: 663–697.

Christoffersen, P., B. Feunou, and Y. Jeon. 2015. "Option Valuation With Observable Volatility and Jump Dynamics." *Journal of Banking and Finance* 61: S101–S120.

Corsi, F. 2009. "A Simple Approximate Long-Memory Model of Realized Volatility." *Journal of Financial Econometrics* 7: 174–196.

Corsi, F., N. Fusari, and D. La Vecchia. 2013. "Realizing Smiles: Options Pricing With Realized Volatility." *Journal of Financial Economics* 107, no. 2: 284–304.

Deng, G., C. McCann, and O. Wang. 2012. "Are VIX Futures ETPs Effective Hedges?" *Journal of Index Investing* 3: 35–48.

Du, D., and D. Luo. 2019. "The Pricing of Jump Propagation: Evidence From Spot and Options Markets." *Management Science* 67: 2360–2387.

Dumas, B., J. Fleming, and R. Whaley. 1998. "Implied Volatility Functions: Empirical Tests." *Journal of Finance* 53: 2059–2106.

Eraker, B. 2004. "Do Stock Prices and Volatility Jump? Reconciling Evidence From Spot and Option Prices." *Journal of Finance* 59: 1367–1404.

Eraker, B., and Y. Wu. 2017. "Explaining the Negative Returns to Volatility Claims: An Equilibrium Approach." *Journal of Financial Economics* 125, no. 1: 72–98.

Feunou, B., and C. Okou. 2019. "Good Volatility, Bad Volatility, and Option Pricing." *Journal of Financial and Quantitative Analysis* 54: 695–727.

Gehricke, S., and J. Zhang. 2018. "Modeling VXX." *Journal of Futures Markets* 38: 958–976.

Gehricke, S., and J. Zhang. 2020a. "The Implied Volatility Smirk in the VXX Options Market." *Applied Economics* 52, no. 8: 769–788.

Gehricke, S., and J. Zhang. 2020b. "Modeling VXX Under Jump Diffusion With Stochastic Long-Term Mean Under Jump Diffusion With Stochastic Long-Term Mean." *Journal of Futures Markets* 40, no. 10: 1508–1534.

Goard, J., and M. Mazur. 2013. "Stochastic Volatility Models and the Pricing of VIX Options." *Mathematical Finance* 23, no. 3: 439–458.

Grasselli, M., A. Mazzoran, and A. Pallavicini. 2023. "A General Framework for a Joint Calibration of VIX and VXX Options." *Annals of Operations Research* 336: 3–26.

Grasselli, M., and L. Wagalath. 2020. "VIX Versus VXX: A Joint Analytical Framework." *International Journal of Theoretical and Applied Finance* 23, no. 5: 1–39.

Guo, Z., Z. Huang, and C. Tong. 2024. "Pricing VIX Futures and Options With Good and Bad Volatility of Volatility." *Journal of Futures Markets* 44: 1832–1847.

Hansen, P., and Z. Huang. 2016. "Exponential GARCH Modeling With Realized Measures of Volatility." *Journal of Business & Economic Statistics* 34: 269–287.

Hansen, P., Z. Huang, and H. Shek. 2012. "Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility." *Journal of Applied Econometrics* 27, no. 6: 877–906.

Hansen, P., and A. Lunde. 2005. "A Realized Variance for the Whole Day Based on Intermittent High-Frequency Data." *Journal of Financial Econometrics* 3: 525–554.

High-Frequency VIX Index. 2023. [data-set] Bloomberg Terminal, and FirstRate data. https://firstratedata.com/.

Huang, J., and L. Wu. 2004. "Specification Analysis of Option Pricing Models Based on Time-Changed Lévy Processes." *Journal of Finance, LIX* 3: 1405–1439.

Jiang, G., G. Qiao, F. Ma, and L. Wang. 2022. "Directly Pricing VIX Futures With Observable Dynamic Jumps Based on High-Frequency VIX." *Journal of Futures Markets* 42: 1518–1548.

Jing, B., S. Li, and Y. Ma. 2020. "Pricing VIX Options With Volatility Clustering." *Journal of Futures Markets* 40: 928–944.

Lin, W., and J. Zhang. 2022. "Pricing VXX Options by Modeling VIX Directly." *Journal of Futures Markets* 42: 888–922.

Lu, S. 2023. "Joint Calibration of VIX and VXX Options: Does Volatility Clustering Matter?" *European Journal of Finance*: 1–32.

Luo, X., J. Zhang, and W. Zhang. 2019. "Instantaneous Sqaured VIX and VIX Derivatives." *Journal of Futures Markets* 39: 1193–1213.

Mencia, J., and E. Sentana. 2013. "Valuation of VIX Derivatives." *Journal of Financial Economics* 108: 367–391.

Newey, W., and K. West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55: 703–708.

Ornthanalai, C. 2014. "Lévy Jump Risk: Evidence From Options and Returns." *Journal of Financial Economics* 112: 69–90.

Pan, J. 2002. "The Jump-Risk Premia Implicit in Options: Evidence From an Integrated Time-Series Study." *Joural of Financial Economics* 63: 3–50.

Pan, Z., Y. Wang, and L. Liu. 2021. "Realized Bipower Variation, Jump Components, and Option Valuation." *Journal of Futures Markets* 41: 1933–1958.

Park, Y.-H. 2016. "The Effects of Asymmetric Volatility and Jumps on the Pricing of VIX Derivatives." *Journal of Econometrics* 192: 313–328.

Qiao, G., and G. Jiang. 2023. "VIX Futures Pricing Based on High-Frequency VIX: A Hybrid Appraoch Combining Svr With Parametric Models." *Journal of Futures Markets* 43: 1238–1260.

Shephard, N., and K. Sheppard. 2010. "Realising the Future: Forecasting With High-Frequency-Based Volatility (HEAVY) Models." *Journal of Applied Econometrics* 25: 197–231.

Tan, X., C. Wang, W. Lin, et al., 2021. "The Term Structure of the VXX Option Smirk: Pricing VXX Option With a Two-Factor Model and Asymmetric Jumps." *Journal of Futures Markets* 41: 439–457.

Tong, C., and Z. Huang. 2021. "Pricing VIX Options With Realized Volatility." *Journal of Futures Markets* 41: 1180–1200.

Tong, C., Z. Huang, and T. Wang. 2022. "Do VIX Futures Contribute to the Valuation of VIX Options." *Journal of Futures Markets* 42: 1644–1664.

VXX Price and Trade Volume. 2024. [data-set] FirstRate data. https://firstratedata.com/.

Wang, T., S. Cheng, F. Yin, and M. Yu. 2022. "Overnight Volatility, Realized Volatility, and Option Pricing." *Journal of Futures Markets* 42: 1264–1283.

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: 135–156.

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: 458–477.

Yoon, J., X. Ruan, and J. Zhang. 2022. "Information Content of the Implied Volatility Smirk of VIX and VXX on SPX Options." Working Paper. University of Otago.

Yuan, P. 2022. "Time-Varying Skew in VIX Derivatives Pricing." *Management Science* 68: 7761–7791.

Zhou, Z., W. Xu, and A. Rubtsov. 2024. "Joint Calibration of S&P 500 and VIX Options Under Local Stochastic Volatility Models." *International Journal of Finance and Economics* 29: 273–310.

Appendix A

Implied Daily VXX Return Dynamics is a Martingale

On the basis of the implied VXX return dynamics in (14) and (15), we have

$$E_t^{\mathbb{Q}}\left[\exp\left(R_{t+1}^{\mathrm{VXX}}\right)\right] = E_t^{\mathbb{Q}}\left[\exp\begin{pmatrix} r + A(1,29) - A(1,30) + B(1,29)h_{z,t+1} - B(1,30)h_{z,t} \\ +\sum_{i=1}^{p} D_i(1,29)y_{t+2-i} - \sum_{i=1}^{p} D_i(1,30)y_{t+1-i} \end{pmatrix}\right]. \tag{A1}$$

Applying recursive relationships in (9), we have

$$\begin{aligned} \sum_{i=1}^{p} D_i(1,29)y_{t+2-i} - \sum_{i=1}^{p} D_i(1,30)y_{t+1-i} &= D_1(1,29)y_{t+1} + \sum_{i=1}^{p-1} D_{i+1}(1,29)y_{t+1-i} - \sum_{i=1}^{p} D_i(1,30)y_{t+1-i} \\ &= D_1(1,29)\left(\beta_0 + \sum_{i=1}^{p}\beta_i y_{t+1-i} + z_{t+1}\right) + \sum_{i=1}^{p-1} D_{i+1}(1,29)y_{t+1-i} - \sum_{i=1}^{p} D_i(1,30)y_{t+1-i} \\ &= D_1(1,29)\beta_0 + D_1(1,29)z_{t+1}. \end{aligned} \tag{A2}$$

Using recursive relationships in (7) and (8) and substituting (A2) into (A1), we have

$$\begin{aligned} E_t^{\mathbb{Q}}\left[\exp\left(R_{t+1}^{\mathrm{VXX}}\right)\right] = \exp&\begin{pmatrix} r - B(1,29)(\omega - a\sigma) + \frac{1}{2}\ln(1 - 2B(1,29)a\sigma) \\ -B(1,29)[b + a + a\sigma((\gamma^*)^2 - \gamma^2)]h_{z,t} \\ -\dfrac{\frac{1}{2}(D_1(1,29))^2 - B(1,29)a\sigma(D_1(1,29))^2(1-\rho^2) + 2B(1,29)a\sigma\gamma^*(B(1,29)a\sigma\gamma^* - D_1(1,29)\rho)}{1 - 2B(1,29)a\sigma}\cdot h_{z,t} \end{pmatrix} \\ &\times E_t^{\mathbb{Q}}[\exp(B(1,29)h_{z,t+1} + D_1(1,29)\sqrt{h_{z,t}}\varepsilon_{1,t+1})]. \end{aligned} \tag{A3}$$

Next, we obtain

$$\begin{aligned} &E_t^{\mathbb{Q}}[\exp(B(1,29)h_{z,t+1} + D_1(1,29)\sqrt{h_{z,t}}\varepsilon_{1,t+1})] \\ &= E_t^{\mathbb{Q}}\left[\exp\begin{pmatrix} B(1,29)\omega + B(1,29)bh_{z,t} + D_1(1,29)\sqrt{h_{z,t}}\varepsilon_{1,t+1} \\ +B(1,29)a\left[h_{z,t} + \sigma((\gamma^*)^2 - \gamma^2)h_{z,t} + \sigma\left[(\varepsilon_{2,t+1} - \gamma^*\sqrt{h_{z,t}})^2 - (1 + (\gamma^*)^2 h_{z,t})\right]\right] \end{pmatrix}\right] \\ &= \exp(B(1,29)\omega + B(1,29)bh_{z,t} + B(1,29)ah_{z,t} + B(1,29)a\sigma((\gamma^*)^2 - \gamma^2)h_{z,t} - B(1,29)a\sigma) \\ &\quad\times E_t^{\mathbb{Q}}\left[\exp\left(B(1,29)a\sigma\varepsilon_{2,t+1}^2 - 2B(1,29)a\sigma\gamma^*\sqrt{h_{z,t}}\varepsilon_{2,t+1} + D_1(1,29)\sqrt{h_{z,t}}\varepsilon_{1,t+1}\right)\right] \\ &= \exp\begin{pmatrix} B(1,29)\omega + B(1,29)bh_{z,t} + B(1,29)ah_{z,t} + B(1,29)a\sigma((\gamma^*)^2 - \gamma^2)h_{z,t} - B(1,29)a\sigma \\ -\frac{1}{2}\ln(1 - 2B(1,29)a\sigma) + \dfrac{\frac{1}{2}(D_1(1,29))^2 - B(1,29)a\sigma(D_1(1,29))^2(1-\rho^2) + 2B(1,29)a\sigma\gamma^*(B(1,29)a\sigma\gamma^* - D_1(1,29)\rho)}{1 - 2B(1,29)a\sigma}\cdot h_{z,t} \end{pmatrix}. \end{aligned} \tag{A4}$$

Finally, substituting (A4) into (A3), we have

$$E_t^{\mathbb{Q}}\left[\exp\left(R_{t+1}^{\mathrm{VXX}}\right)\right] = e^r \tag{A5}$$

Appendix B

MGF of VXX Return: HAR-RV-GARCH

For simplicity, let $H = H(B(1,29))$, $\Lambda = \Lambda(B(1,29), B(1,30))$, which are defined in (18) and (19). On the basis of the daily VXX return dynamics in (16) and (17) implied from the HAR-RV-GARCH model, the one-step ahead conditional MGF of VXX_T at time t under $\mathbb{Q}$ has an exponential form:

$$\begin{aligned} E_t^{\mathbb{Q}}\left[\exp\left(uR_{t+1}^{\mathrm{VXX}} + vh_{z,t+1}\right)\right] &= E_t^{\mathbb{Q}}\left[\exp\begin{pmatrix} u(r + H + \Lambda h_{z,t} + B(1,29)aRV_{t+1} + D_1(1,29)z_{t+1}) \\ +v(\omega + bh_{z,t} + aRV_{t+1}) \end{pmatrix}\right] \\ &= \exp\begin{pmatrix} u(r+H) + v\omega - \sigma(uB(1,29)a + va) + u\Lambda h_{z,t} + vbh_{z,t} \\ (uB(1,29)a + va)(1 + \sigma((\gamma^*)^2 - \gamma^2))h_{z,t} \end{pmatrix} \\ &\quad\times E_t^{\mathbb{Q}}\left[\exp\begin{pmatrix} \sigma(uB(1,29)a + va)\left(\varepsilon_{2,t+1}^2 - 2\gamma^*\sqrt{h_{z,t}}\varepsilon_{2,t+1}\right) \\ +uD_1(1,29)\sqrt{h_{z,t}}\varepsilon_{1,t+1} \end{pmatrix}\right] \end{aligned} \tag{B1}$$

and we have

$$E_t^{\mathbb{Q}}\left[\exp\begin{pmatrix}\sigma(uB(1,29)a+\nu a)\left(\varepsilon_{2,t+1}^2-2\gamma^*\sqrt{h_{z,t}}\varepsilon_{2,t+1}\right)\\ +uD_1(1,29)\sqrt{h_{z,t}}\varepsilon_{1,t+1}\end{pmatrix}\right] = \exp\begin{pmatrix}-\frac{1}{2}\ln(1-2\sigma(uB(1,29)a+\nu a))+\frac{(uD_1(1,29)\rho-2\sigma\gamma^*(uB(1,29)a+\nu a))^2}{2(1-2\sigma(uB(1,29)a+\nu a))}\cdot h_{z,t}\\ +\frac{1}{2}(uD_1(1,29))^2(1-\rho^2)h_{z,t}\end{pmatrix}. \tag{B2}$$

Substituting (B2) into (B1), we obtain

$$\begin{aligned}
&E_t^{\mathbb{Q}}\left[\exp\left(uR_{t+1}^{\text{VXX}}+\nu h_{z,t+1}\right)\right]\\
&= \exp\begin{pmatrix}u(r+H)+\nu\omega-\sigma(uB(1,29)a+\nu a)-\frac{1}{2}\ln(1-2\sigma(uB(1,29)a+\nu a))\\ +u\Lambda h_{z,t}+\nu b h_{z,t}+(uB(1,29)a+\nu a)(1+\sigma((\gamma^*)^2-\gamma^2))h_{z,t}\\ +\frac{(uD_1(1,29)\rho-2\sigma\gamma^*(uB(1,29)a+\nu a))^2}{2(1-2\sigma(uB(1,29)a+\nu a))}\cdot h_{z,t}+\frac{1}{2}(uD_1(1,29))^2(1-\rho^2)h_{z,t}\end{pmatrix}\\
&= \exp(E(u,\nu)h_{z,t}+G(u,\nu))
\end{aligned} \tag{B3}$$

with

$$\begin{aligned}
E(u,\nu) &= u\Lambda+\nu b+(uB(1,29)a+\nu a)(1+\sigma((\gamma^*)^2-\gamma^2))\\
&\quad+\frac{(uD_1(1,29)\rho-2\sigma\gamma^*(uB(1,29)a+\nu a))^2}{2(1-2\sigma(uB(1,29)a+\nu a))}+\frac{1}{2}(uD_1(1,29))^2(1-\rho^2),\\
G(u,\nu) &= u(r+H)+\nu\omega-\sigma(uB(1,29)a+\nu a)-\frac{1}{2}\ln(1-2\sigma(uB(1,29)a+\nu a)).
\end{aligned}$$

We conjecture that the τ-step ahead MGF has the following form:

$$\Psi_{t,t+\tau}(u)=E_t^{\mathbb{Q}}\left[\exp\left(u\sum_{j=1}^{\tau}R_{t+j}^{\text{VXX}}\right)\right]=\exp(Q(u,\tau)h_{z,t}+R(u,\tau)) \tag{B4}$$

and we have

$$\begin{aligned}
E_t^{\mathbb{Q}}\left[\exp\left(u\sum_{j=1}^{\tau+1}R_{t+j}^{\text{VXX}}\right)\right] &= E_t^{\mathbb{Q}}\left[E_{t+1}^{\mathbb{Q}}\left[\exp\left(u\sum_{j=1}^{\tau+1}R_{t+j}^{\text{VXX}}\right)\right]\right]\\
&= E_t^{\mathbb{Q}}\left[\exp\left(uR_{t+1}^{\text{VXX}}\right)E_{t+1}^{\mathbb{Q}}\left[\exp\left(u\sum_{j=2}^{\tau+1}R_{t+j}^{\text{VXX}}\right)\right]\right]\\
&= E_t^{\mathbb{Q}}\left[\exp\left(uR_{t+1}^{\text{VXX}}+Q(u,\tau)h_{z,t+1}+R(u,\tau)\right)\right]\\
&= \exp(E(u,Q(u,\tau))h_{z,t}+G(u,Q(u,\tau))+R(u,\tau)),
\end{aligned} \tag{B5}$$

which yields

$$\begin{aligned}
Q(u,\tau+1) &= E(u,Q(u,\tau)),\\
R(u,\tau+1) &= G(u,Q(u,\tau))+R(u,\tau)
\end{aligned}$$

subject to the following initial conditions:

$$\begin{aligned}
Q(u,1) &= E(u,0),\\
R(u,1) &= G(u,0).
\end{aligned}$$

Appendix C

MGF of VXX Return: HAR-GARCH

On the basis of the daily VXX return dynamics in (20) and (21) implied from the HAR-GARCH model, the conditional MGF of VXX_T at time t has an exponential affine form:

$$
\begin{aligned}
f^{\text{VXX}}(u;t,T) &= E_t^{\mathbb{Q}}\left[e^{u\ln \text{VXX}_T}\right] = e^{u\ln \text{VXX}_t}\cdot E_t^{\mathbb{Q}}\left[\exp\left(u\sum_{j=1}^{\tau} R_{t+j}^{\text{VXX}}\right)\right] \\
&= e^{u\ln \text{VXX}_t}\cdot \exp(Q(u;t,T)h_{z,t+1} + R(u;t,T)).
\end{aligned}
$$

Using iterated expectation, we have

$$
\begin{aligned}
f^{\text{VXX}}(u;t,T) &= E_t^{\mathbb{Q}}\left[E_{t+1}^{\mathbb{Q}}\left[e^{u\ln \text{VXX}_T}\right]\right] = E_t^{\mathbb{Q}}\left[f^{\text{VXX}}(u;t+1,T)\right] \\
&= E_t^{\mathbb{Q}}\left[e^{u\ln \text{VXX}_{t+1}}\cdot \exp(Q(u;t+1,T)h_{z,t+2} + R(u;t+1,T))\right] \\
&= e^{u\ln \text{VXX}_t}\cdot E_t^{\mathbb{Q}}\left[\exp\left(uR_{t+1}^{\text{VXX}} + Q(u;t+1,T)h_{z,t+2} + R(u;t+1,T)\right)\right].
\end{aligned}
$$

Since the values of $Q(u;t,T)$ and $R(u;t,T)$ only depend on $\tau = T - t$, we simplify the notations to $Q(u,\tau)$ and $R(u,\tau)$, respectively. Therefore, we have the τ-step ahead conditional MGF of VXX returns given by

$$
\begin{aligned}
\Psi_{t,t+\tau}(u) &= E_t^{\mathbb{Q}}\left[\exp\left(u\sum_{j=1}^{\tau} R_{t+j}^{\text{VXX}}\right)\right] = \exp(Q(u,\tau)h_{z,t+1} + R(u,\tau)) \\
&= E_t^{\mathbb{Q}}\left[\exp\left(uR_{t+1}^{\text{VXX}} + Q(u,\tau-1)h_{z,t+2} + R(u,\tau-1)\right)\right] \\
&= \exp\left(\begin{array}{l}
u(r+H) + Q(u,\tau-1)\omega + R(u,\tau-1) \\
-\frac{1}{2}\ln(1 - 2[uB(1,29)a + Q(u,\tau-1)a]) \\
+u\Lambda h_{z,t+1} + Q(u,\tau-1)bh_{z,t+1} \\
+\dfrac{\frac{1}{2}(uD_1(1,29))^2 + \gamma^*(\gamma^* - 2uD_1(1,29))[uB(1,29)a + Q(u,\tau-1)a]}{1 - 2[uB(1,29)a + Q(u,\tau-1)a]}h_{z,t+1}
\end{array}\right),
\end{aligned}
$$

which yields the following recursive relationships:

$$
\begin{aligned}
Q(u,\tau) &= u\Lambda + Q(u,\tau-1)b \\
&\quad + \frac{\frac{1}{2}(uD_1(1,29))^2 + \gamma^*(\gamma^* - 2uD_1(1,29))[uB(1,29)a + Q(u,\tau-1)a]}{1 - 2[uB(1,29)a + Q(u,\tau-1)a]}, \\
R(u,\tau) &= u(r+H) + Q(u,\tau-1)\omega + R(u,\tau-1) \\
&\quad - \frac{1}{2}\ln(1 - 2[uB(1,29)a + Q(u,\tau-1)a])
\end{aligned}
$$

subject to the following terminal conditions at time T: $Q(u,0) = R(u,0) = 0$. H and Λ functions are defined in (22) and (23).