Check for updates

Received: 25 February 2022 | Accepted: 9 July 2022

DOI: 10.1002/fut.22372

RESEARCH ARTICLE

FUTURES MARKETS WILEY

Forecasting realized volatility: New evidence from time-varying jumps in VIX

Anupam Dutta[1] | **Debojyoti Das**[2]

[1]School of Accounting and Finance, University of Vaasa, Vaasa, Finland

[2]Finance & Accounting Area, Indian Institute of Management Bangalore, Bengaluru, Karnataka, India

Correspondence
Anupam Dutta, School of Accounting and Finance, University of Vaasa, Vaasa, Finland.
Email: adutta@uwasa.fi

Abstract

Given that jumps in the implied volatility index (VIX) lead to rapid changes in the level of volatility, they may contain significant predictive information for the realized variance (RV) of stock returns. Against this backdrop, the present study proposes to extend the heterogeneous autoregressive (HAR) model using the information content of time-varying jumps occurring in VIX. We find that jumps in VIX have positive impacts on the RV of S&P 500 index and that the proposed HAR-RV approach generates more accurate volatility forecasts than do the existing HAR-RV type models. Importantly, these results hold for short-, medium-, and long-term volatility components.

KEYWORDS

HAR model, jump intensity, jump size, VIX, volatility forecasts, volatility jumps

JEL CLASSIFICATION

G1

1 | INTRODUCTION

Modeling stock market volatility is an important step to precisely measure the risk of investor portfolio. An accurate measure of true volatility is of utmost importance to investors for developing appropriate hedging strategies to mitigate potential risks associated with their investments. Therefore, over the years investors, policymakers, and researchers have been in search of proper estimates of stock price volatility. This, however, remains a complex task given the time-varying features of equity return volatility.

Given the importance of finding accurate measures of stock market volatility, this paper aims to extend the prior literature by investigating whether time-varying jumps in implied volatility index (VIX), which occur as a consequence of terrorist attacks, natural disasters, recession, or political violence, provide additional information which could be useful for the prediction of future realized volatility (RV) of S&P 500 index. Detecting the presence of time-varying jumps in VIX is important as such jumps lead to rapid changes in the level of volatility and due to volatility persistence, a long-lasting impact on the distribution of S&P 500 returns. In addition, the occurrence of these jumps raises the probability of extreme movements in the underlying asset returns (Aït-Sahalia & Hurd, 2015; Bandi & Reno, 2016). Some recent studies (Wang et al., 2022; Yin et al., 2021) also argue that jumps in VIX are important for capturing the shocks of large and rare events, which will reinforce the pricing accuracy when the

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2022 The Authors. The *Journal of Futures Markets* published by Wiley Periodicals LLC.

market is highly volatile.[1] Given that volatility jumps tend to increase the risk in financial markets, such jumps could have significant predictive content for the realized variance of asset returns.

In particular, our analysis attempts to answer several important empirical questions for the US stock market. First, do time-varying jumps occurring in VIX provide incremental forecasting information for the equity return volatility? Second, are volatility jumps better predictors of future RV compared to the VIX index? Third, does introducing the jump intensity (JI) of VIX to the heterogeneous autoregressive or HAR-type models improve their forecasting power?

The literature on forecasting stock price volatility is relatively extensive. Since the inception of GARCH-type models during the 1980s, a strand of literature has investigated the predictive role of such approaches when analyzing the volatility of asset returns. These studies document that different versions of GARCH models provide good in-sample estimates of stock market volatility (see Bera & Higgins, 1993; Bollerslev, 1987; Bollerslev et al., 1992; Bollerslev et al., 1994; Glosten et al., 1993; Nelson, 1991). However, numerous studies reveal that GARCH-type models are incapable of producing good forecasts of stock price volatility (Akgiray, 1989; Brailsford & Faff, 1996; Figlewski, 1997; Franses & Van Dijk, 1996; Kat & Heynen, 1994). All these papers document that the R^2 statistics do not increase markedly when regressing RV on forecast volatility. Moreover, Dimson and Marsh (1990) find that although data snooping has the ability to provide good in-sample estimates, but it fails to do so for out-of-sample forecasting. In addition, Nelson (1992) employs theoretical methods to evidence that GARCH models generate good in-sample estimates for high-frequency data, but out-of-sample forecasting still remains poor.

Another line of literature provides empirical evidence that the informational content of VIX plays a key role in forecasting stock market volatility. Dennis et al. (2006), for instance, document that implied volatilities successfully predict future RV of the S&P 100 equity index and 50 large US firms. In addition, Carr and Wu (2006) show that VIX has better predictive power than traditional GARCH approaches. Yu et al. (2010) also find similar evidence. A recent study by Kambouroudis and McMillan (2016) reveals that the information content of VIX index plays a crucial role in forecasting the RV of global capital markets. More recently, Pati et al. (2018) find that while VIX index is a biased forecast, it possesses relevant information in explaining the future RV. However, mixed views exist here as well. Day and Lewis (1992) show that VIX fails to offer additional content that could provide better forecasts for stock market volatility. Canina and Figlewski (1993) also document that implied volatilities appear to be poor forecasts of volatility and that simple historical volatilities have better predictive power than implied volatilities. The findings of Becker et al. (2006) conclude the same as well.

Recently, the HAR-type models have received considerable attention for predicting the RV of global stock markets. The HAR approach considers separation of RV into short-, medium-, and long-term volatility components. Numerous studies (Andersen et al., 2007; Andersen et al., 2011; Forsberg & Ghysels, 2007; Giot & Laurent, 2007; Ma et al., 2014) find that when forecasting the future RV, the HAR-RV process outperforms the GARCH-type, SV-type, VAR-RV, MIDAS-RV, and ARFIMA-RV approaches. Considering its success in forecasting future RV, this model has been extended in several ways. For instance, Andersen et al. (2007) consider decomposition of RV into continuous sample path and discontinuous jump components to propose HAR-RV-J and HAR-RV-CJ models. The authors evidence a significant improvement in model accuracy. Moreover, Andersen et al. (2011) make use of overnight return variance to develop the HAR-RV-CJN model and show that the HAR-RV-CJN model yields better volatility forecasts than do the GARCH and HAR-RV approaches. In addition, Çelik & Ergin (2014) also conclude that the HAR-RV model outperforms GARCH process when predicting the volatility of the Turkish equity market.

However, the results of earlier studies are somewhat conflicting. For example, Andersen et al. (2007), Forsberg and Ghysels (2007), Giot and Laurent (2007), and Busch et al. (2011) document either a negative or no impact of jumps on the RV of asset returns. Corsi et al. (2010), on the other hand, show that such effect is positive and that considering the information on volatility jumps improves the forecast accuracy.

This empirical research joins the existing literature to examine whether the information on time-varying jumps in VIX improves the predictive power of HAR-type models. In doing so, our study makes at least two contributions. First,

[1]The works of Wang et al. (2022) and Yin et al. (2021) model the VIX index using the heterogeneous autoregressive (HAR) structure and conclude that such a process gives a sufficient description of the underlying VIX series when pricing VIX derivatives. These studies also document that while applying the HAR process, the incorporation of jumps in VIX generates reliable pricing performances under various market circumstances.

while numerous studies (Andersen et al., 2007; Busch et al., 2011; Corsi et al., 2010; Forsberg & Ghysels, 2007; Giot & Laurent, 2007) consider splitting up RV into its continuous sample path (C) and jump (J) components to use them as separate regressors when forecasting future RV, this paper uses the JI of VIX index to propose a new HAR-RV model for forecasting the RV of S&P 500 returns. Based on our knowledge, this is first study to make this attempt. Since time-varying jumps arriving in VIX lead to a significant increment in the volatility of the underlying asset, such jumps could contain predictive information for future RV of the US stock market.

Our second contribution comes from the empirical findings. We show that jumps in VIX have a positive and highly significant impact on the RV of S&P 500 index and that these jumps carry additional information beyond what is contained in the existing HAR-RV models. Our results thus extend the works of Giot and Laurent (2007) and Busch et al. (2011) in that these studies evidence that the information content of VIX improves the forecasting power of the baseline HAR-RV models, while we show that replacing VIX with its jump component leads to a further improvement in the accuracy of volatility forecasts. Our results also demonstrate that during periods of market stress, the extended models generate more accurate out-of-sample forecasts than do the existing HAR approaches. Given that jumps in VIX allow volatility to rapidly increase which provides some signals of market crash, the proposed HAR-RV models that include the information on jumps improve the accuracy of volatility forecasts amid the stress periods. Hence, the results of this study can be exploited to understand the economics of stock markets further.

Notably, we employ the GARCH-jump process, proposed by Chan and Maheu (2002), to measure the JI for the VIX index and then use this information to extend the HAR-RV process to predict the RV of S&P 500 index. Note that we have extended the HAR-RV models proposed by Andersen et al. (2007) and Corsi (2009). In total, we have considered eight HAR-RV type models to compare the forecast accuracy.

This study has important implications to market participants. For instance, jumps in VIX capture both good and bad surprises which may arrive with different rates and sizes and hence investors may react differently to them (Park, 2016). In addition, the conditional jump size (JS) distribution impacts not only the first- and second-order conditional moments of the VIX index, but also the flexible higher-order conditional moments (i.e., skewness and kurtosis) of the same. Therefore, the information on JSs and JI provides a good depiction of implied volatilities which could be useful for understanding the US equity market volatility amid the stress periods (Wang et al., 2022; Yin et al., 2021). In sum, jumps in the VIX could offer additional information which helps market participants predict the RV more precisely.

The rest of the paper proceeds as follows. Section 2 outlines the methods. The data used in this empirical research are described in Section 3. We discuss the findings in Section 4, while the last section concludes our paper.

2 | METHODOLOGY

2.1 | GARCH-jump approach

The GARCH-jump process has achieved immense popularity in recent finance literature (Chiang et al., 2019; Dutta et al., 2021; Gronwald, 2019; Zhou et al., 2019). In this study, we consider the following form for this process:

$$\Delta VIX_t = a + \mu \Delta VIX_{t-1} + \epsilon_t, \tag{1}$$

where ΔVIX_t denotes the first-order difference of VIX index and ϵ_t is the error divided into two components:

$$\epsilon_t = \epsilon_{1t} + \epsilon_{2t}, \tag{2}$$

with ϵ_{1t} having the following form:

$$\epsilon_{1t} = \sqrt{h_t} z_t, z_t \sim NID(0, 1) h_t = \omega + \alpha \epsilon_{1t-1}^2 + \beta h_{t-1}. \tag{3}$$

Additionally, ϵ_{2t} refers to a jump innovation:

$$\epsilon_{2t} = \sum_{l=1}^{n_t} U_{tl} - \theta\lambda_t, \tag{4}$$

where U_{tl} indicates the JS which is normally distributed with a mean θ and a variance ϑ^2. $\sum_{l=1}^{n_t} U_{tl}$ presents the jump factor, and n_t denotes the number of jumps, which follows a Poisson distribution as follows:

$$P(n_t = j|I_{t-1}) = \frac{e^{-\lambda_t}\lambda_t^j}{j!}, j = 0,1,2, \ldots \tag{5}$$

where j indicates the number of jumps at time t and λ_t is the autoregressive conditional jump intensity (ARJI) defined as:

$$\lambda_t = \lambda_0 + \rho\lambda_{t-1} + \gamma\xi_{t-1}, \tag{6}$$

where $\lambda_t > 0$, $\lambda_0 > 0$, $\rho > 0$, and $\gamma > 0$. Time-varying jumps are said to occur in VIX, if the parameters ρ, γ are statistically significant. Note that ρ measures jump persistence, whereas ξ_{t-1} refers to the innovation to λ_{t-1}, which is modeled as:

$$\xi_{t-1} = E[n_{t-1}|I_{t-1}] - \lambda_{t-1}. \tag{7}$$

In Equation (7), $E[n_{t-1}|I_{t-1}]$ is the ex post assessment of the expected number of jumps that occurred from $t-2$ to $t-1$, which can be obtained by

$$E[n_{t-1}|I_{t-1}] = \sum_{j=0}^{\infty} jP(n_{t-1} = j|I_{t-1}) \tag{8}$$

After integrating out all of the jumps during a one-unit interval, the conditional probability density function of ΔVIX_t can be expressed as:

$$P(\Delta VIX_t|I_{t-1}) = \sum_{j=0}^{\infty} f(\Delta VIX_t|n_t = j, I_{t-1})P(n_t = j|I_{t-1}), \tag{9}$$

where $f(\Delta VIX_t|n_t = j, I_{t-1})$ represents the conditional density of ΔVIX_t given j jumps occurring up to time $t-1$ and follows a Gaussian distribution. Then the likelihood function can be constructed as:

$$f(\Delta VIX_t|n_t = j, I_{t-1}) = (2\pi(h_t + j\vartheta^2))^{-1/2} exp \times \left(-\frac{(\Delta VIX_t - \mathrm{a} - \mu\Delta VIX_{t-1} - \theta j)^2}{2(h_t + j\vartheta^2)}\right). \tag{10}$$

Hence, the log-likelihood is given by

$$L(\Omega) = \sum_{t=1}^{T} \log f(\Delta VIX_t|I_{t-1}; \Omega), \tag{11}$$

where $\Omega = (\mathrm{a}, \mu, \omega, \alpha, \beta, \theta, \vartheta, \lambda_0, \rho, \gamma)$.

Note that if the conditional JI is stationary, that is, ($|\rho| < 1$), then the unconditional JI will be equal to

$$E(\lambda_t) = \frac{\lambda_0}{1 - \rho}. \tag{12}$$

Moreover, following Chan and Maheu (2002), the multiperiod forecasts of the expected number of future jumps can be expressed as

$$E(\lambda_{t+i}|I_{t-1}) = \lambda_0(1 + \rho + \cdots + \rho^{i-1}) + \rho^i\lambda_t, i \geq 1. \tag{13}$$

In case, when $i = 1$, $E(\lambda_{t+i}|I_{t-1}) = \lambda_0 + \rho\lambda_t$.

2.2 | HAR-RV models

Let us define the RV for day t as follows:

$$RV_t = \sum_{j=1}^{M} r_{t,j}^2;\ t = 1,2, \ldots, T \tag{14}$$

where $r_{t,j}$ is the logarithmic return for period j of day t, M indicates the number of intraday observations at time t and T refers to the number of periods in the sample. Based on the theory of Barndorff-Nielsen and Shephard (2004), RV can be satisfied when $M \to \infty$:

$$RV_t \to \int_0^t \sigma^2(s)ds + \sum_{0<s\leq t} \kappa^2(s), \tag{15}$$

where $\int_0^t \sigma^2(s)ds$ is the continuous component, while $\sum_{0<s\leq t} \kappa^2$ indicates the jump component. When $M \to \infty$, the continuous component is approximately equal to the realized bi-power variation (BV) defined as

$$BV_t = \frac{\pi}{2}\frac{M}{M-2}\sum_{j=3}^{M} |r_{t,j}||r_{t,j-2}|;\ t = 1, 2, \ldots, T \tag{16}$$

Then in argument with Corsi et al. (2010), the HAR-RV model is specified as

$$\text{HAR-RV}: RV_{t,t+h} = \tau_0 + \tau_d RV_t + \tau_w RV_{t-5,t} + \tau_m RV_{t-22,t} + \varepsilon_t \tag{17}$$

where h takes the values 1, 5 and 22 for daily, weekly and monthly RV models, respectively and

$$RV_{t_1,t_2} = \frac{1}{t_2 - t_1}\sum_{t=t_1+1}^{t_2} RV_t. \tag{18}$$

Hence, in the HAR-RV model, we decompose the RV of S&P 500 index into short-, medium-, and long-term volatility components.

As mentioned earlier, Andersen et al. (2007) consider decomposition of RV into continuous sample path (C) and discontinuous jump components (J) to propose the following model:

$$RV_{t,t+h} = \tau_0 + \tau_{cd}C_t + \tau_{cw}C_{t-5,t} + \tau_{cm}C_{t-22,t} + \tau_j J_t + \varepsilon_t, \tag{19}$$

where,

$$C_t = RV_t - J_t;\ t = 1,2, \ldots, T \tag{20}$$

with

$$J_t = I_{\{Z_t > \phi_{1-\alpha}\}}(RV_t - BV_t);\ t = 1,2, \ldots, T. \tag{21}$$

In Equation (21), $I_{\{E\}}$ denotes the indicator for the event E, $\phi_{1-\alpha}$ is the $100(1-\alpha)\%$ point in the standard normal distribution, and α refers to the significance level.

Note that the Z-ratio test statistic, proposed by Barndorff-Nielsen and Shephard (2006), is defined as

$$Z_t = \sqrt{M}\frac{(RV_t - BV_t)RV_t^{-1}}{\sqrt{\left(\frac{\pi^2}{4} + \pi - 5\right)\max\left\{1, \frac{TQ_t}{BV_t^2}\right\}}}, \tag{22}$$

with TQ_t referring to the realized Tripower Quarticity, which is defined as follows:

$$TQ_t = u_{4/3}^{-3}\frac{M^2}{M-4}\sum_{j=5}^{M}|r_{t,j}|^{4/3}|r_{t,j-k-1}|^{4/3}|r_{t,j-2k-2}|^{4/3}, \tag{23}$$

where $u_{4/3} = 2^{2/3}\Gamma(7/6)/\Gamma(1/2)$. In the absence of jumps, $Z_t \to N(0,1)$ as $M \to \infty$, and large positive values indicate the presence of jumps during period t.

Note that a jump is said to be significant when $Z_t > \phi_{1-\alpha}$. Following Andersen et al. (2007) and Busch et al. (2011), we choose $\alpha = 0.01$ to detect jumps.

Moreover, Busch et al. (2011) extend model (19) as follows:

$$\text{HAR-CJ}: RV_{t,t+h} = \tau_0 + \tau C_t + \tau_{cw}C_{t-5,t} + \tau_{cm}C_{t-22,t} + \tau_{jd}J_t + \tau_{jw}J_{t-5,t} + \tau_{jm}J_{t-22,t} + \varepsilon_t. \tag{24}$$

It is worth mentioning that Busch et al. (2011) use the informational content of VIX index to extend models (17) and (24) as follows[2]:

$$\text{HAR-RV-IV}: RV_{t,t+h} = \tau_0 + \tau_d RV_t + \tau_w RV_{t-5,t} + \tau_m RV_{t-22,t} + \tau_{VIX}VIX_t + \varepsilon_t, \tag{25}$$

$$\text{HAR-CJ-IV}: RV_{t,t+h} = \tau_0 + \tau C_t + \tau_{cw}C_{t-5,t} + \tau_{cm}C_{t-22,t} + \tau_{jd}J_t + \tau_{jw}J_{t-5,t} + \tau_{jm}J_{t-22,t} + \tau_{VIX}VIX_t + \varepsilon_t. \tag{26}$$

Busch et al. (2011) find evidence that VIX provides incremental information about future RV and HAR-RV-IV or HAR-CJ-IV process outperforms the baseline regression approach. In this paper, we propose to replace VIX index by its JI component (i.e., λ_t) to arrive at the following specifications:

$$\text{HAR-RV-JI}: RV_{t,t+h} = \tau_0 + \tau_d RV_t + \tau_w RV_{t-5,t} + \tau_m RV_{t-22,t} + \tau_\lambda \lambda_t + \varepsilon_t, \tag{27}$$

$$\text{HAR-CJ-JI}: RV_{t,t+h} = \tau_0 + \tau C_t + \tau_{cw}C_{t-5,t} + \tau_{cm}C_{t-22,t} + \tau_{jd}J_t + \tau_{jw}J_{t-5,t} + \tau_{jm}J_{t-22,t} + \tau_\lambda \lambda_t + \varepsilon_t. \tag{28}$$

Our objective is to compare the proposed models (27) and (28) with those presented in Equations (17) and (24)–(26). Hence, six models are considered in our empirical analyses.[3]

In line with Busch et al. (2011), we consider using the following models to generate forecasts for the continuous component (C) and jump component (J), respectively:

$$C_{t,t+h} = \tau_0 + \tau_{cd}C_t + \tau_{cw}C_{t-5,t} + \tau_{cm}C_{t-22,t} + \varepsilon_t, \tag{29}$$

$$J_{t,t+h} = \tau_0 + \tau_{jd}J_t + \tau_{jw}J_{t-5,t} + \tau_{jm}J_{t-22,t} + \varepsilon_t, \tag{30}$$

where $C_{t,t+h}(J_{t,t+h})$ replaces $RV_{t,t+h}$ as regressand compared to Equation (17).

[2]Unlike Fernandes et al. (2014), we consider ΔVIX instead of VIX as the VIX index, based on the standard unit root tests, appears to be weekly stationary at levels. For the difference series, these tests, however, suggest that we can strongly reject the presence of a unit root.

[3]In all our regression specifications, nonoverlapping windows are used. Busch et al. (2011), for instance, argue that the use of overlapping windows may cause serial correlation in the error term. Christensen and Prabhala (1998) also document that employing overlapping windows may lead to erroneous inferences in a setting involving both implied volatility and RV. Although using overlapping windows does not necessarily undermine the parameter estimates, an adjustment still needs to be made for the accurate estimation of standard errors.

2.3 | Forecast evaluation

2.3.1 | Root mean squared error (RMSE)

In addition to finding the best-fitted models, it is also important to identify which of them appears to be the superlative approach for out-of-sample volatility prediction. To address this issue, we consider the RMSE statistic in our analysis. The RMSE statistic is defined as

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(RV_t - \widehat{RV_t})^2}, \tag{31}$$

where T indicates the number of forecast data points, while RV_t and $\widehat{RV_t}$ are the actual volatility and predicted volatility for day t, respectively. Note that as a measure for true volatility, we consider the 5-min intra-day squared returns for the S&P 500 index.

Moreover, to assess the null hypothesis of no difference in accuracy associated with different models, we also consider the application of Diebold and Mariano (DM) test (2002).

To explain this test, let $e_{it} = RV_t - \widehat{RV_t}$ $(i = 1,2)$ be the forecast errors. Now let $d_t = f(e_{1t}) - f(e_{2t})$, where $f(.)$ is a function of forecast errors. Then the null hypothesis can be specified as:

$$H_0 : E(d_t) = 0.$$

DM (2002) show that $\bar{d}$ has an approximate asymptotic variance given as

$$Var(\bar{d}) \approx n^{-1}\left[\delta_0 + 2\sum_{l=1}^{q-1}\delta_l\right] \tag{32}$$

with δ_l being the lth autocovariance of d_t, estimated as:

$$\hat{\delta} = n^{-1}\sum_{t=l+1}^{n}(d_t - \bar{d})(d_{t-k} - \bar{d}). \tag{33}$$

Then the DM test statistic is defined as:

$$\text{DM} = (\widehat{Var(\bar{d})})^{-1/2}\bar{d}. \tag{34}$$

Under H_0, the DM has an asymptotic standard normal distribution.

2.3.2 | Mincer–Zarnowitz (MZ) regression

The MZ (1969) regression approach is employed to investigate whether the forecasts involving JI deliver additional information beyond what is contained in the existing HAR-RV models. The baseline MZ regression model is given by:

$$RV_t = \varphi_0 + \varphi_1\widehat{RV_t} + \epsilon_t \tag{35}$$

where RV_t and $\widehat{RV_t}$ refer to the true volatility and volatility forecast for day t, respectively. We then calculate the coefficient of determination (i.e., R^2) to compare the prediction performances.

2.3.3 | Forecast encompassing tests

The forecast encompassing testing procedure, proposed by Chong and Hendry (1986), is used to examine whether a forecast obtained from a competing model outperforms a base model forecast. In particular, if the competing model does not provide further information, then the base model forecast wins the battle. That is, the competing model fails to encompass the traditional approach. To apply this test, we extend Equation (9) as follows:

$$RV_t = \varphi_0 + \varphi_1 \widehat{RV}_{1,t} + \varphi_2 \widehat{RV}_{2,t} + \epsilon_t. \tag{36}$$

In Equation (36), subscripts 1 and 2 indicate the forecast models with 1 denoting the base model and 2 being the competing model. The null hypothesis for forecast encompassing test is that model 1 outperforms model 2 suggesting that φ_2 is equal to zero. On the other hand, if φ_2 is significant and positive, then model 2 contains information that model 1 does not, such that model 2 is not encompassed by model 1 (Chong & Hendry, 1986).

3 | DATA

We utilize the daily observations of the US VIX and S&P 500 indexes. The sample period ranges from January 1, 2000 to June 30, 2020. The information on VIX is retrieved from the Thomson Reuters DataStream database, while the 5-min intra-day squared returns for the S&P 500 index are taken from Oxford-Man Institute's Realized Library version 0.3. Note that the in-sample estimation period spans from January 1, 2000 to June 30, 2018, whereas the out-of-sample period ranges from July 1, 2018 to June 30, 2020. It is also worth mentioning that the VIX index is multiplied by $1/\sqrt{252}$ and considered in variance form.

Table 1 presents the descriptive statistics for realized and implied volatility measures. It is observed that the daily, weekly and monthly means of RV, continuous and jump components appear to be equal, though the standard deviations tend to vary. In addition, all the variables are right skewed and exhibit kurtosis higher than three. Therefore, these indexes do not follow Gaussian distributions. In particular, the values of skewness and kurtosis indicate the presence of sharp picks and fat tail distributions.

The VIX index is depicted in Figure 1. A number of spikes or jumps are observed over the sample period considered. These jumps can be attributed to economic recessions, geo-political uncertainty or epidemic outbreak. For example, the most prominent jumps are detected amid the 2008 global financial crisis and the ongoing COVID-19 pandemic periods. Besides, the jump identified at the beginning of 2011 could be the consequences of the Libyan war due to which oil prices experience a significant downturn and as a result, uncertainty in global financial market tends to increase.

4 | EMPIRICAL RESULTS

4.1 | Estimates of the GARCH-jump process

Table 2 shows the findings of the GARCH-jump approach. The findings suggest that the GARCH parameters (α, β) are strongly significant at 1% level indicating the evidence of volatility clustering. We further notice that $\alpha < \beta$ suggesting the significance of long-term persistence. Basher et al. (2016) and Ahmad et al. (2018) also report similar findings.

The results also reveal the existence of jumps in VIX and confirm that such jumps tend to vary over time given that the intensity parameters ρ, γ are significant at conventional levels. The parameter ρ, assessing the persistence in the conditional JI, appears to be 0.9716. Moreover, the JI parameters ρ, γ assume positive values, which would justify the choice of ARJI specification when detecting time-dependent jumps in VIX. In addition, the positivity of these coefficients indicates that λ_t, the JI, is influenced by both lagged JI (λ_{t-1}) and lagged intensity residuals (ξ_{t-1}).

Our findings are in line with Baldeaux and Badran (2014), Bandi and Renò (2016) and Ait-Sahalia and Hurd (2015). The papers cited above also document the presence of time-dependent jumps in the US implied VIX. However, these papers mainly use stochastic volatility or diffusion models to detect the occurrence of jumps in VIX. In addition, the authors of these cited studies do not use the information content of time-varying JI and JSs associated with the VIX index while predicting the RV of S&P 500 returns.

We report the summary statistics of JI for the VIX index in Table A1. These numbers suggest that the mean intensity amounts to 0.1424 and that there do not exist any negative values as the intensity (λ_t) lies between 0.0043

TABLE 1 Summary statistics

	Mean	SD	Maximum	Minimum	Skewness	Kurtosis
RV-daily	0.000112	0.000269	0.007748	0.000001	10.65	191.53
RV-weekly	0.000112	0.000230	0.003420	0.000003	7.31	73.96
RV-monthly	0.000112	0.000198	0.002110	0.000005	5.74	43.57
C-daily	0.000090	0.000224	0.006018	0.000001	10.47	170.51
C-weekly	0.000090	0.000195	0.002826	0.000002	7.63	77.87
C-monthly	0.000090	0.000168	0.001752	0.000004	5.90	44.55
J-daily	0.000022	0.000074	0.002025	0.000000	12.76	248.76
J-weekly	0.000022	0.000046	0.000712	0.000000	6.78	69.22
J-monthly	0.000022	0.000035	0.000388	0.000000	4.63	33.25
VIX	0.000208	0.595731	13.86241	−9.331000	3.10	139.76

Note: The summary statistics for RV, continuous sample path components (*C*), jump components (*J*) and VIX index. Here, we consider the first differences for the VIX index.

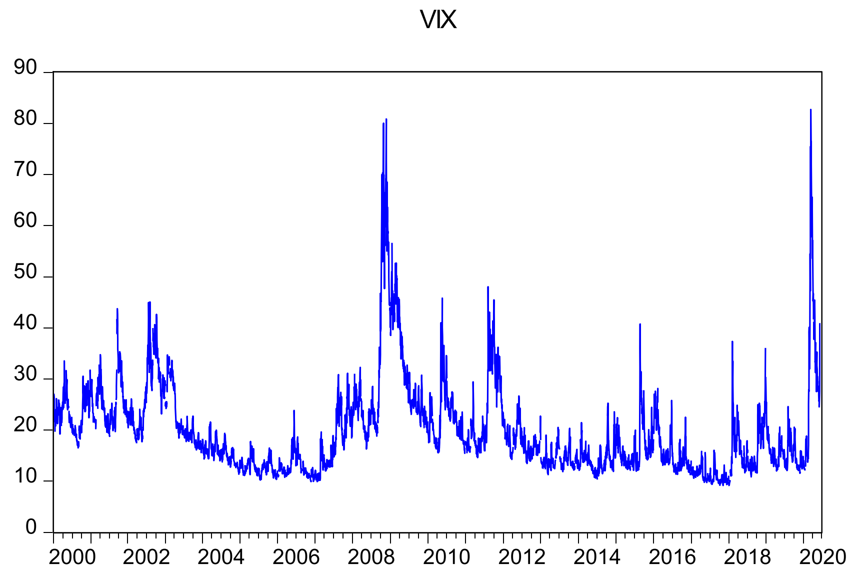


FIGURE 1 Implied volatility index for the period 2000–2020. This figure exhibits the US equity market implied volatility index (VIX) which is derived from the S&P 500 options for the 30 days following the measurement date. As shown in the diagram, the sample period contains important recessions including the 2008 global financial crises and the COVID-19 pandemic. [Color figure can be viewed at wileyonlinelibrary.com]

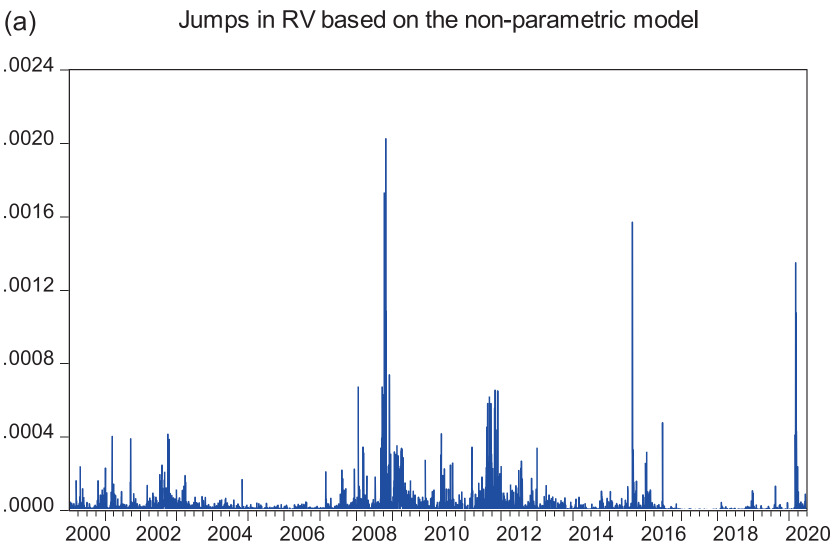
and 1.1172. This simply indicates that the intensity parameter satisfies the non-negativity condition. We also notice that λ_t attains its maximum value during the COVID-19 pandemic period. This jump seems signaling the global stock market crashes due to the novel coronavirus disease which is widely known as the "Coronavirus Crash." Prior literature (see Aït-Sahalia & Hurd, 2015; Bandi & Reno, 2016) also argues that the occurrence of jumps in implied volatilities tends to increase the probability of extreme movements in the underlying asset returns, providing an indication of bear markets.

Next, Figure 2 depicts the jump component from the nonparametric model (a) along with the time-varying JI for the VIX index (b). While both diagrams demonstrate some uniformity in regard to the occurrence of jumps, the latter one seems to reveal more information about the volatility dynamics of the S&P 500 index. For example, each of these graphs indicates the presence of jumps during the 2008 global financial crisis, the 2014 oil price decline and the ongoing COVID-19 pandemic periods. However, we also detect clusters of jumps in VIX following the September 11 attacks, which is not the case for the jump component of RV. Besides, the significance of the Libyan war taking place in 2011 or the 2018 cryptocurrency crash also triggers a number of jumps in the VIX index, which is unlikely as well when looking at the jumps in RV. Note that in each of these cases (i.e., the September 11 attacks, the Libyan war and the 2018 cryptocurrency crash), the US stock prices experience a significant fall and spikes are, therefore, apparent in the implied VIX. Hence, it can be concluded that jumps in implied volatility appear to be a better indicator of market conditions than that in RV.

TABLE 2 Results of GARCH-jump process

Variables	Estimates	Standard errors	*t* statistics	*p* values
a	−0.00510***	0.0012	−4.12	0.00
μ	−0.09364***	0.0150	−6.28	0.00
ω	0.00045***	0.0001	6.56	0.00
α	0.22680***	0.0181	12.50	0.00
β	0.69720***	0.0207	33.58	0.00
θ	0.24702***	0.0364	6.78	0.00
ϑ^2	0.34451***	0.0235	14.64	0.00
λ_0	0.00421	0.0028	1.53	0.12
ρ	0.97160***	0.0195	49.82	0.00
γ	0.11155**	0.0445	2.51	0.02
Log-likelihood	1529.50			

Note: The results for the GARCH-jump process applied to the VIX index. *** and ** indicate statistical significance at 1% and 5% levels, respectively.



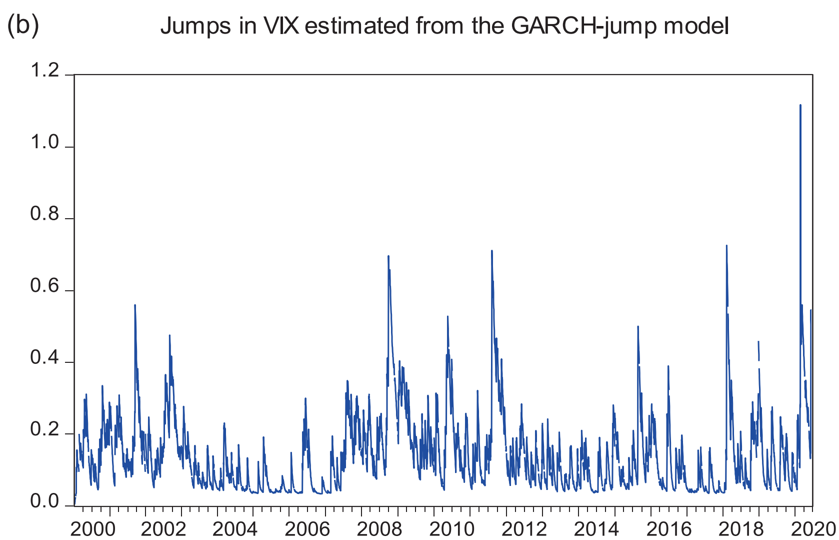


FIGURE 2 Jumps in RV and jump intensities for VIX. This figure depicts the jump component of RV from the non-parametric model (a) and the time-varying jump intensities for the VIX index (b). For (a), the jump component of realized volatility (RV) is defined in Equation (21). For (b), the time-varying jump intensities are calculated from Equation (4). [Color figure can be viewed at wileyonlinelibrary.com]

4.2 | Results of HAR-RV models

The results of HAR-RV type models are shown in Tables 3–5. The findings presented in Table 3 refer to the estimates of 1-day future volatility. Looking at the baseline HAR-RV model, it is evident that the coefficients of short-, medium- and long-term volatility components (i.e., τ_d, τ_w, τ_m) are all significant at 1% level implying that the US stock market volatility has a long memory. Similar results are also observed for the HAR-RV-JI approach. For instance, the coefficients are still significant at 1% level. More importantly, the JI parameter is also highly significant and positive. In addition, a significant increment is also observed in the R^2 statistic. For the HAR-RV-IV model, we observe that the effect of VIX index is positive and significant. However, the size of its impact is smaller than that of λ_{t-1}, though the R^2 statistic increases slightly (from 54.1% to 54.4%). This finding is in line with Giot and Laurent (2007) and Busch et al. (2011) who also document the importance of including VIX in the HAR-RV type models. Next, moving to the estimates HAR-CJ, HAR-CJ-JI, and HAR-CJ-IV models, we report several interesting findings. First, the R^2 statistics improve markedly. In particular, the HAR-CJ-JI approach yields the highest R^2 value (57.7%) followed by the HAR-CJ-IV model (54.9%). Second, the coefficients of different jump components appear to be negative in majority of the cases. While this result is consistent with earlier studies (see Andersen et al., 2007; Busch et al., 2011; Corsi et al., 2010; Giot & Laurent, 2007), the economic

TABLE 3 Daily realized volatility HAR models

Models	HAR-RV	HAR-RV-JI	HAR-RV-IV	HAR-CJ	HAR-CJ-JI	HAR-CJ-IV
τ_0	1.1×10^{-5}*** (2.3×10^{-6})	-7.4×10^{-6}* (4.2×10^{-6})	9.0×10^{-6}*** (3.1×10^{-6})	1.0×10^{-5}*** $(3.\ 2 \times 10^{-6})$	-1.0×10^{-5}** (4.0×10^{-6})	1.0×10^{-5}*** (τ_d)
τ_d	0.2732*** (0.0176)	0.2716*** (0.0175)	0.1914*** (0.0178)			
τ_w	0.3285*** (0.0232)	0.3748*** (0.0297)	0.4046*** (0.0230)			
τ_m	0.2265*** (0.0258)	0.1991*** (0.0262)	0.2173*** (0.0250)			
τ_{cd}				0.3790*** (0.0231)	0.3738*** (0.0230)	0.3013*** (0.0227)
τ_{cw}				0.6684*** (0.0407)	0.6377*** (0.0409)	0.7684*** (0.0396)
τ_{cm}				−0.1102*** (0.0418)	−0.1031** (0.0417)	−0.1185*** (0.0404)
τ_{jd}				−0.1244*** (0.0433)	−0.1175*** (0.0432)	−0.2292*** (0.0422)
τ_{jw}				−0.7072*** (0.1074)	−0.7743*** (0.1075)	−0.6206*** (0.1038)
τ_{jm}				1.6123*** (0.1610)	1.3811*** (0.1645)	1.5886*** (0.1554)
τ_{IV}			8.3×10^{-5}** (4.1×10^{-5})			8.6×10^{-5}*** (5.1×10^{-6})
τ_λ		1.8×10^{-4}*** (3.2×10^{-5})			2.0×10^{-4}*** (3.3×10^{-5})	
R^2 (%)	54.1	56.8	54.4	54.3	57.7	54.9
HET test	0.21	0.67	0.76	0.17	0.58	0.89

Note: The in-sample estimates for the daily HAR-RV models. The in-sample estimation period spans from January 1, 2000 to June 30, 2018, yielding a total of 4654 daily observations. Six different models are considered in our empirical analyses. These models are presented in Equations (17), (24)–(28). Values in parentheses indicate the standard errors. We also provide the estimates of R^2 statistics. The values shown in the last row refer to the p values for the F test of no heteroscedasticity. ***, **, and * imply that the null hypothesis is rejected at 1%, 5%, and 10% significance levels, respectively.

TABLE 4 Weekly realized volatility HAR models

Models	HAR-RV	HAR-RV-JI	HAR-RV-IV	HAR-CJ	HAR-CJ-JI	HAR-CJ-IV
τ_0	2.1×10^{-5}*** (2.0 × 10^{-6})	-5.0×10^{-6} (3.1 × 10^{-6})	1.7×10^{-5}*** (2.4 ×10^{-6})	-1.0×10^{-5} (3.0 × 10^{-6})	2.0×10^{-5}*** (2.1 × 10^{-6})	
τ_d	0.1754*** (0.0140)	0.1733*** (0.0138)	0.1253*** (0.0142)			
τ_w	0.2997*** (0.0231)	0.2525*** (0.0235)	0.3580*** (0.0231)			
τ_m	0.3612*** (0.0205)	0.3252*** (0.0207)	0.3556*** (0.0202)			
τ_{cd}				0.2455*** (0.0183)	0.2399*** (0.0182)	0.1978*** (0.0182)
τ_{cw}				0.5234*** (0.0322)	0.4904*** (0.0322)	0.5847*** (0.0318)
τ_{cm}				−0.0440 (0.0331)	−0.0365 (0.0329)	−0.0491 (0.0324)
τ_{jd}				−0.1135*** (0.0343)	−0.1061*** (0.0340)	−0.1777*** (0.0339)
τ_{jw}				−0.6754*** (0.0851)	−0.7474*** (0.0849)	−0.6223*** (0.0834)
τ_{jm}				2.1577*** (0.1275)	1.9094*** (0.1298)	2.1432*** (0.1249)
τ_{IV}			5.1×10^{-5}*** (4.0 × 10^{-6})			5.3×10^{-5}*** (4.0 × 10^{-6})
τ_λ		2.3×10^{-4}*** (2.5 × 10^{-5})			2.2×10^{-4}*** (2.5 × 10^{-5})	
R^2 (%)	59.1	61.7	60.1	62.2	64.9	62.7
HET test	0.19	0.55	0.47	0.13	0.32	0.73

Note: The in-sample estimates for the weekly HAR-RV models. The in-sample estimation period spans from January 1, 2000 to June 30, 2018, yielding a total of 4654 daily observations. The weekly measures are the scaled sums of the corresponding daily measures. Six different models are considered in our empirical analyses. These models are presented in Equations (17), (24)–(28). Values in parentheses indicate the standard errors. We also provide the estimates of R^2 statistics. The values shown in the last row refer to the *p* values for the *F* test of no heteroscedasticity. ***, **, and * imply that the null hypothesis is rejected at 1%, 5%, and 10% significance levels, respectively.

intuition would suggest an upsurge in volatility following a jump in the price process (Corsi et al., 2010). Third, the coefficient of JI is significantly positive and its size is higher than that of the VIX parameter. Forth, the coefficients of long-term volatility components (i.e., τ_{cm}) are negative for the HAR-CJ models, whereas the corresponding estimates are positive for the HAR-RV models.

Next, Tables 4 and 5 report the estimates from the weekly and monthly volatility forecast models. These results are almost in line with those shown in Table 3 with few exceptions. First, for both weekly and monthly volatility models, the coefficients of long-term volatility components are found to be insignificant in case of HAR-CJ models, while the corresponding estimates are significantly positive for the HAR-RV models. Second, the R^2 statistics improve markedly for the weekly models, although this is not the case for the monthly volatility forecasts. Third, the daily jump components are mostly insignificant for the monthly volatility models. Note that Tables 3–5 also report the *p* values for the *F* test of no heteroscedasticity. The findings suggest that the null hypothesis of no heteroscedasticity is never rejected at the conventional levels.

Overall, the results indicate that the information on volatility jumps could be useful when predicting the RV of the US equity market. In the following sections, we examine whether the HAR-CJ-IV model also provides the best out-of-sample forecasts.

TABLE 5 Monthly realized volatility HAR models

Models	HAR-RV	HAR-RV-JI	HAR-RV-IV	HAR-CJ	HAR-CJ-JI	HAR-CJ-IV
τ_0	3.0×10^{-5}*** (2.0×10^{-6})	1.1×10^{-5}*** (4.2×10^{-6})	3.0×10^{-5}*** (2.0×10^{-6})	3.2×10^{-5}*** (2.0×10^{-6})	1.3×10^{-5}*** (3.0×10^{-6})	2.7×10^{-5}*** (2.0×10^{-6})
τ_d	0.1022*** (0.0130)	0.1003*** (0.0129)	0.0711*** (0.0134)			
τ_w	0.3087*** (0.0215)	0.2678*** (0.0219)	0.3449*** (0.0217)			
τ_m	0.2973*** (0.0191)	0.2661*** (0.0193)	0.2938*** (0.0190)			
τ_{cd}				0.1512*** (0.0174)	0.1474*** (0.0174)	0.1226*** (0.0176)
τ_{cw}				0.3102*** (0.0307)	0.2882*** (0.0308)	0.3470*** (0.0307)
τ_{cm}				0.0349 (0.0316)	0.0400 (0.0314)	0.0319 (0.0313)
τ_{jd}				−0.0530 (0.0327)	−0.0481 (0.0326)	−0.0916*** (0.0327)
τ_{jw}				0.1999*** (0.0810)	0.1513* (0.0811)	0.2311*** (0.0805)
τ_{jm}				1.5809*** (0.1214)	1.4154*** (0.1241)	1.5722*** (0.1204)
τ_{IV}			3.2×10^{-5}*** (3.0×10^{-6})			3.2×10^{-5}*** (3.0×10^{-6})
τ_λ		2.0×10^{-4}*** (2.3×10^{-5})			1.4×10^{-4}*** (2.4×10^{-5})	
R^2 (%)	54.4	56.1	55.0	55.8	57.2	56.1
HET test	0.62	0.42	0.49	0.53	0.71	0.59

Note: The in-sample estimates for the monthly HAR-RV models. The in-sample estimation period spans from January 1, 2000 to June 30, 2018, yielding a total of 4654 daily observations. The monthly measures are the scaled sums of the corresponding daily measures. Six different models are considered in our empirical analyses. These models are presented in Equations (17), (24)–(28). Values in parentheses indicate the standard errors. We also provide the estimates of R^2 statistics. The values shown in the last row refer to the p values for the F test of no heteroscedasticity. ***, **, and * imply that the null hypothesis is rejected at 1%, 5%, and 10% significance levels, respectively.

4.3 | Forecasting evaluation based on RMSE and DM test

Table 6 displays the values of the RMSE statistic and the DM test results. It is evident that for both daily and weekly volatility models, the HAR-RV-JI approach appears to have the lowest RMSE statistics. For the monthly models, on the other hand, we report the lowest RMSE statistic for the HAR-CJ-JI specification. These results confirm that our proposed models produce more accurate volatility forecasts than do the existing HAR-RV approaches.

The DM test, however, fails to reject the null hypothesis of no difference in accuracy for the daily HAR-RV models. When looking at the medium-term volatility component models, our findings suggest that the HAR-RV-JI approach outperforms all other specifications. Finally, for the long-term volatility component, the HAR-CJ-JI specification seems to be outclassing its counterparts. Given that the results do not lead to any specific conclusions for the daily volatility models, the estimates of MZ regressions and forecast encompassing tests seem important for identifying the best forecast models for the short-term volatility component.

TABLE 6 RMSE values and DM test results

	Daily HAR models		Weekly HAR models		Monthly HAR models	
	RMSE	DM statistic	RMSE	DM statistic	RMSE	DM statistic
HAR-RV	0.000254	0.9193	0.000273	2.8302***	0.000302	3.9432***
HAR-RV-JI	**0.000253**		**0.000268**		0.000296	1.6576***
HAR-RV-IV	0.000268	0.7921	0.000269	1.2931*	0.000301	2.1482**
HAR-CJ	0.000264	1.0073	0.000285	1.6162*	0.000295	3.8859***
HAR-CJ-JI	0.000263	0.9786	0.000282	1.3938*	**0.000292**	
HAR-CJ-IV	0.000276	1.1843	0.000281	1.2895*	0.000295	1.6497***

Note: The values of RMSE statistic and the Diebold-Mariano (DM) test results. The in-sample estimation period spans from January 1, 2000 to June 30, 2018, whereas the out-of-sample period ranges from July 1, 2018 to June 30, 2020. The numbers in bold indicate the lowest values for the RMSE statistic. We employ the DM test to assess the null hypothesis of no difference in accuracy for different models. ***, **, and * imply that the null hypothesis of the DM test is rejected at the 1%, 5%, and 10% significance levels, respectively.

TABLE 7 Results of Mincer and Zarnowitz (MZ) regression models

	R^2 (%)		
Models	**Daily**	**Weekly**	**Monthly**
HAR-RV	61.1	44.4	13.5
HAR-RV-JI	63.8	48.4	14.9
HAR-RV-IV	56.8	46.0	14.4
HAR-CJ	58.5	40.7	14.5
HAR-CJ-JI	58.8	41.6	15.5
HAR-CJ-IV	55.9	42.6	15.3

Note: This table reports the R^2 (%) values from the Mincer and Zarnowitz (MZ) regression model. This process assumes the following form: $RV_t = \varphi_0 + \varphi_1 \hat{RV}_t + \epsilon_t$, where RV_t and $\hat{RV}_t$ refer to the true volatility and volatility forecast for day t, respectively. The sum of 5-min intra-day squared returns is used as a proxy for the true volatility. The in-sample estimation period spans from January 1, 2000 to June 30, 2018, whereas the out-of-sample period ranges from July 1, 2018 to June 30, 2020. These results are obtained for short-term (daily), medium-term (weekly), and long-term (monthly) volatility components.

4.4 | Results of Mincer and Zarnowitz (MZ) regression approach

Table 7 reports the R^2 (%) values from the MZ regressions. For the daily volatility models, the HAR-RV-JI process outperforms other approaches as it generates higher R^2 value (63.8%) than its counterparts. Moving to the weekly volatility models, the HAR-RV-JI approach also provides superior forecasts ($R^2 = 48.4\%$) than the remaining specifications.

When looking at the monthly volatility models, the HAR-CJ-JI yields the best forecasts with the highest R^2 value. This finding is important given that it extends the empirical work of Busch et al. (2011) who show that the information content of VIX improves the forecasting power of the baseline HAR-RV models when forecasting long-term (monthly) volatility component. However, we show that considering the information on time-varying jumps occurring in VIX leads to a further improvement in the accuracy of volatility forecasts.

On the whole, the best forecast models, based on the MZ regression results, appear to be those approaches which consider the information on time-varying jumps occurring in the VIX index. These findings simply indicate that the inclusion of JI factor tends to increase the R^2 value of the MZ regressions. Therefore, the evidence of jumps in VIX adds to the information content of the HAR-RV model when forecasting stock market variability.

4.5 | Results of forecast encompassing tests

Tables 8–10 exhibit the findings of forecast encompassing tests. Note that based on the RMSE statistics and MZ regression results, we choose the HAR-RV-JI model as the baseline approach for both daily and weekly cases. For the

TABLE 8 Results of forecast encompassing test (daily HAR models)

	$RV_t = \varphi_0 + \varphi_1\widehat{RV}_{1,t} + \varphi_2\widehat{RV}_{2,t} + \epsilon_t$				
	HAR-RV-JI versus HAR-RV	HAR-RV-JI versus HAR-RV-IV	HAR-RV-JI versus HAR-CJ	HAR-RV-JI versus HAR-CJ-JI	HAR-RV-JI versus HAR-CJ-IV
φ_0	−0.000016 (0.000013)	−0.000011 (0.000012)	−0.000011 (0.000013)	−0.000011 (0.000013)	−0.000011 (0.000012)
φ_1	2.2003*** (0.7414)	1.1345*** (0.1431)	1.1329*** (0.1781)	1.1821*** (0.1931)	1.1198*** (0.1302)
φ_2	−1.1087 (0.7309)	−0.0549 (0.1319)	−0.0496 (0.1550)	−0.0936 (0.1659)	−0.0367 (0.1072)
R^2(%)	61.8	61.6	61.7	61.7	61.6
F statistic	400.05***	397.27***	397.18***	397.45***	397.20***

Note: The forecast encompassing test results for the daily HAR models. To obtain these outcomes, we estimate the following regression process: $RV_t = \varphi_0 + \varphi_1\widehat{RV}_{1,t} + \varphi_2\widehat{RV}_{2,t} + \epsilon_t$, where the subscripts 1 and 2 indicate the forecast models with 1 denoting the base model and 2 being the competing model. Based on the RMSE values and MZ regression results, we choose the HAR-RV-JI model as the baseline approach. The in-sample estimation period spans from January 1, 2000 to June 30, 2018, whereas the out-of-sample period ranges from July 1, 2018 to June 30, 2020. Values in parentheses denote standard errors. *** indicates statistical significance at 1% level.

TABLE 9 Results of forecast encompassing test (weekly HAR models)

	$RV_t = \varphi_0 + \varphi_1\widehat{RV}_{1,t} + \varphi_2\widehat{RV}_{2,t} + \epsilon_t$				
	HAR-RV-JI versus HAR-RV	HAR-RV-JI versus HAR-RV-IV	HAR-RV-JI versus HAR-CJ	HAR-RV-JI versus HAR-CJ-JI	HAR-RV-JI versus HAR-CJ-IV
φ_0	−0.000008 (0.000014)	0.000010 (0.000014)	−0.000001 (0.000013)	0.000001 (0.000013)	0.000009 (0.000014)
φ_1	3.4651*** (0.6199)	0.4775** (0.2269)	1.3720*** (0.1880)	1.4265*** (0.2126)	0.8836*** (0.1587)
φ_2	−2.4482*** (0.6070)	−0.4782** (0.2146)	−0.3780** (0.1719)	−0.4287** (0.1956)	0.0819 (0.1405)
R^2(%)	47.7	46.5	46.5	46.3	46.0
F-statistic	223.17***	212.71***	212.59***	212.56***	208.43***

Note: The forecast encompassing test results for the weekly HAR models. To obtain these outcomes, we estimate the following regression process: $RV_t = \varphi_0 + \varphi_1\widehat{RV}_{1,t} + \varphi_2\widehat{RV}_{2,t} + \epsilon_t$, where the subscripts 1 and 2 indicate the forecast models with 1 denoting the base model and 2 being the competing model. Based on the RMSE values and MZ regression results, we choose the HAR-RV-JI model as the baseline approach. The in-sample estimation period spans from January 1, 2000 to June 30, 2018, whereas the out-of-sample period ranges from July 1, 2018 to June 30, 2020. Values in parentheses denote standard errors. *** and ** indicate statistical significance at 1% and 5% levels, respectively.

monthly forecast models, the HAR-CJ-JI model is used as the baseline approach. The results of Table 8 reveal that all the φ_1 coefficients (parameters of HAR-RV-JI process) are positive and statistically significant at 1% level. However, the parameters for competing models appear to be insignificant suggesting that the HAR-RV-JI process encompasses its counterparts. This finding is consistent with what is found from the MZ regression approach. Hence the results demonstrate that the JI of VIX has sufficient explanatory power for equity return volatility.

Moving to Table 9, we find that the forecasts obtained from the HAR-RV-JI process encompass those achieved from other approaches. This finding is also in line with that shown in Tables 6 and 7. It is also observed that although the coefficients (φ_2) of HAR-RV, HAR-RV-IV, HAR-CJ, and HAR-CJ-JI models are significant, these are negative and much smaller than φ_1. Therefore, the economic significance of φ_2 is debatable.

Next, the results of Table 10 confirm that the HAR-CJ-JI model provides more accurate forecasts for the long-term volatility component in comparison to other models. This finding is consistent with that obtained from the DM tests and the MZ regression results showing that the HAR-CJ-JI process outperforms its counterparts.

In sum, the HAR-RV-JI and HAR-CJ-JI models, based on the results of different forecast evaluation methods, are found to surpass all the existing HAR-RV specifications while predicting future RV of the S&P 500 index. We

TABLE 10 Results of forecast encompassing test (monthly HAR models)

	$RV_t = \varphi_0 + \varphi_1\widehat{RV}_{1,t} + \varphi_2\widehat{RV}_{2,t} + \epsilon_t$				
	HAR-CJ-JI versus HAR-RV	**HAR-CJ-JI versus HAR-RV-JI**	**HAR-CJ-JI versus HAR-RV-IV**	**HAR-CJ-JI versus HAR-CJ**	**HAR-CJ-JI versus HAR-CJ-IV**
φ_0	0.000085*** (0.000014)	0.000077*** (0.000015)	0.000075*** (0.000014)	0.000047*** (0.000015)	0.000075*** (0.000015)
φ_1	2.4656*** (0.4534)	0.8315* (0.4453)	0.8749*** (0.3190)	6.4117*** (1.0594)	0.9132** (0.4218)
φ_2	−1.7990*** (0.4249)	−0.2608 (0.4271)	−0.2927 (0.2932)	−5.8272*** (1.0537)	0.1550 (0.3630)
R^2(%)	18.6	15.6	15.7	20.6	15.5
F statistic	53.96***	43.57***	43.94***	61.45***	43.46***

Note: The forecast encompassing test results for the monthly HAR models. To obtain these outcomes, we estimate the following regression process: $RV_t = \varphi_0 + \varphi_1\widehat{RV}_{1,t} + \varphi_2\widehat{RV}_{2,t} + \epsilon_t$ where the subscripts 1 and 2 indicate the forecast models with 1 denoting the base model and 2 being the competing model. Based on the RMSE values and MZ regression results, we choose the HAR-CJ-JI model as the baseline approach. The in-sample estimation period spans from January 1, 2000 to June 30, 2018, whereas the out-of-sample period ranges from July 1, 2018 to June 30, 2020. Values in parentheses denote standard errors. ***, **, and * indicate statistical significance at 1%, 5%, and 10% levels, respectively.

can thus conclude that the evidence of time-dependent jumps in VIX provides additional information that is not captured by other HAR-RV type models. Hence both HAR-RV-JI and HAR-CJ-JI approaches provide incremental information for future RV of the US stock market. In other words, including JI in the HAR-RV model improves its goodness of fit.

4.6 | Subsample analysis

In this section, we examine how the extended HAR-RV models perform during the low and high volatility periods. In doing so, we consider the following cases:

Case I. The in-sample estimation period spans from January 1, 2000 to January 31, 2019, whereas the out-of-sample period ranges from February 1, 2019 to January 31, 2020.

Case II. The in-sample estimation period spans from January 1, 2000 to January 31, 2020, whereas the out-of-sample period ranges from February 1, 2020 to June 30, 2020.

In Case I and Case II, we investigate the predictive power of the proposed models for the low and high volatility regimes, respectively. A close look at Figures 1 and 2 helps us to choose these periods. Each of these diagrams indicates that the volatility levels remain relatively low throughout the year 2019, while several spikes or jumps are observed during the COVID-19 period starting from the second month of 2020.

The results of our subsample analyses are presented in Table 11. Panel A shows the results of Case I, whereas Panel B does the same for Case II. Now the findings, which are based on the MZ regression models, reveal several interesting facts.[4] First, for the low volatility regime, the HAR-RV-IV approach provides the best predictions for the daily volatility model as evidenced by the R^2 values. This finding is not in line with our previous analysis, although the results remain the same for the weekly and monthly volatility models. Second, the HAR-RV-JI model outperforms all other specifications during the high volatility periods and this result holds for daily, weekly, and monthly volatility models. While this finding is also inconsistent with our prior investigation, it seems that during the COVID-19 periods, capturing the extreme movements (e.g., jumps) in VIX plays a pivotal role in understanding the volatility dynamics of the S&P 500 index.

[4]We find nearly similar results when employing forecast-encompassing tests.

TABLE 11 Results of subsample analysis using the MZ regression models

Models	R^2 (%) Daily	Weekly	Monthly
Panel A: Case I			
HAR-RV	58.9	40.3	14.0
HAR-RV-JI	63.4	48.2	15.1
HAR-RV-IV	63.7	46.3	15.0
HAR-CJ	59.2	41.9	14.8
HAR-CJ-JI	59.5	43.0	17.9
HAR-CJ-IV	59.6	43.2	15.7
Panel B: Case II			
HAR-RV	53.2	38.9	11.9
HAR-RV-JI	58.6	45.4	15.7
HAR-RV-IV	55.6	42.7	12.6
HAR-CJ	54.0	39.2	12.1
HAR-CJ-JI	55.2	42.6	12.9
HAR-CJ-IV	55.4	42.9	12.8

Note: The R^2 (%) values of the subsample analysis using the Mincer and Zarnowitz (MZ) regression model. This MZ process assumes the following form: $RV_t = \varphi_0 + \varphi_1\widehat{RV_t} + \varepsilon_t$, where RV_t and $\widehat{RV_t}$ refer to the true volatility and volatility forecast for day t, respectively. The sum of 5-min intra-day squared returns is used as a proxy for the true volatility. For Case I, the in-sample estimation period spans from January 1, 2000 to January 31, 2019, whereas the out-of-sample period ranges from February 1, 2019 to January 31, 2020 and for Case II, the in-sample estimation period spans from January 1, 2000 to January 31, 2020, whereas the out-of-sample period ranges from February 1, 2020 to June 30, 2020. These results are obtained for short-term (daily), medium-term (weekly), and long-term (monthly) volatility components.

Overall, these results confirm that jumps represent an important element of risk and that replacing VIX with its jump component leads to a further improvement in the accuracy of volatility forecasts. Besides, we can also conclude that jumps in implied volatility do reveal more information than that in RV.

4.7 | Robustness tests

In this section, we intend to compare the results of our HAR models with respect to GARCH specifications. As the HAR-RV-JI and HAR-CJ-JI models appear to have better predictive power than other HAR approaches, we mainly consider these two models in our robustness analysis. Besides, we also employ two additional HAR models which include both VIX and JI of VIX. These two new models are referred to as HAR-RV-IV-JI and HAR-CJ-IV-JI.

It is also worth mentioning that the GJR-GARCH model of Glosten et al. (1993) has been adopted in our robustness analysis.[5] As the GARCH model has also been extended by including VIX and JI, we have the following specifications: GJR-GARCH, GJR-GARCH-IV, GJR-GARCH-JI, and GJR-GARCH-IV-JI. We now define the GJR-GARCH process as follows:

$$r_t = m + nr_{t-1} + \varepsilon_t, \tag{37}$$

$$\varepsilon_t = e_t\sigma_t, \tag{38}$$

[5]Giot and Laurent (2007) also employ the GJR-GARCH approach arguing that rather parsimonious models seem to perform well. Besides, the AIC, BIC, and likelihood function values also recommend this asymmetric specification.

TABLE 12 Results of MZ regression models for testing robustness

	R^2 (%)		
Models	**Daily**	**Weekly**	**Monthly**
HAR-RV-JI	61.6	48.2	14.9
HAR-RV-IV-JI	62.9	45.9	14.9
HAR-CJ-JI	58.8	41.6	17.6
HAR-CJ-IV-JI	58.9	41.5	15.2
GJR-GARCH	56.3	32.7	12.3
GJR-GARCH-IV	58.6	40.0	14.0
GJR-GARCH-JI	59.4	40.9	14.6
GJR-GARCH-IV-JI	60.1	41.1	14.6

Note: The R^2 (%) values from the Mincer and Zarnowitz (MZ) regression model. The main objective of this presentation is to compare the results of HAR models with respect to GARCH specifications. Besides, we also consider two additional HAR models which include both VIX and jump intensities (JI) of VIX. The MZ process assumes the following form: $RV_t = \varphi_0 + \varphi_1 R\hat{V}_t + \varepsilon_t$, where RV_t and $R\hat{V}_t$ refer to the true volatility and volatility forecast for day t, respectively. The sum of 5-min intra-day squared returns is used as a proxy for the true volatility. The in-sample estimation period spans from January 1, 2000 to June 30, 2018, whereas the out-of-sample period ranges from July 1, 2018 to June 30, 2020. These results are obtained for short-term (daily), medium-term (weekly), and long-term (monthly) volatility components.

where r_t is the daily return for the S&P 500 index, e_t is drawn from the standardized normal distribution and σ_t is defined as[6]:

$$\text{GJR-GARCH}: \sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \delta\varepsilon_{t-1}^2 d_{t-1}, \tag{39}$$

where δ measures the asymmetric effect and d_t refers to a dummy variable, which equals 1 when ε_t is negative. The GJR-GARCH model is then extended as follows:

$$\text{GJR-GARCH-IV}: \sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \delta\varepsilon_{t-1}^2 d_{t-1} + \chi VIX_{t-1}, \tag{40}$$

$$\text{GJR-GARCH-JI}: \sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \delta\varepsilon_{t-1}^2 d_{t-1} + \varsigma\lambda_{t-1}, \tag{41}$$

$$\text{GJR-GARCH-IV-JI}: \sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \delta\varepsilon_{t-1}^2 d_{t-1} + \chi VIX_{t-1} + \varsigma\lambda_{t-1}. \tag{42}$$

To generate the h-day volatility forecasts using the GJR-GARCH models, we adopt an out-of-sample estimation approach following Giot (2005) and Giot and Laurent (2007). In particular, we obtain the volatility forecasts using a fixed rolling window scheme: the estimation period is rolled forward by adding one new daily observation and dropping the most distant observation. Thus, the volatility forecasts are generated for the next day and then for the next 5 days (1 week) and 22 days (1 month). Note that we have an update every 5 days (for the one– and five–steps–ahead forecasts) and every 22 days for a 1-month horizon.

Table 12 reports the R^2 (%) values from the MZ regression model. These results are mostly in line with those exhibited in Table 7. For example, for the weekly volatility models, the HAR-RV-JI approach yields the best forecasts followed by the HAR-RV-IV-JI and HAR-CJ-JI specifications. Moving to the monthly volatility models, we find that the HAR-CJ-JI process still outperforms other models by generating superior forecasts. However, looking at the daily volatility models, we find somewhat inconsistent results. We now observe that the HAR-RV-IV-JI model provides the best predictions followed by the HAR-RV-JI and GJR-GARCH-IV-JI processes.

Overall, our results indicate that including both VIX and JI usually improves the predictive power of the HAR or GARCH models, albeit the HAR-RV-JI and HAR-CJ-JI specifications appear to have the best forecasts for the weekly and monthly volatility components, respectively. These findings are also supported by the forecast

[6]It is noteworthy that the AR(1) model given in Equation (37) is selected on the basis of AIC and BIC values.

TABLE 13 Results of forecast encompassing tests for robustness analysis

	$RV_t = \varphi_0 + \varphi_1\widehat{RV}_{1,t} + \varphi_2\widehat{RV}_{2,t} + \epsilon_t$					
	Daily HAR models HAR-RV-IV-JI versus HAR-RV-JI		Weekly HAR models HAR-RV-JI versus HAR-RV-IV-JI		Monthly HAR models HAR-CJ-JI versus HAR-CJ-IV-JI	
	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
φ_0	0.000001	0.000012	−0.000143***	0.000031	−0.000182	0.000029
φ_1	0.774191***	0.000010	0.732018***	0.061776	0.000181***	0.000004
φ_2	0.236415	0.151462	0.000143	0.000228	0.000044	0.000032
$R^2(\%)$	66.9		67.5		67.4	
F statistic	480.13***		513.96***		512.55***	

Note: The forecast encompassing test results for robustness analysis. To obtain these outcomes, we estimate the following regression process: $RV_t = \varphi_0 + \varphi_1\widehat{RV}_{1,t} + \varphi_2\widehat{RV}_{2,t} + \epsilon_t$, where the subscripts 1 and 2 indicate the forecast models with 1 denoting the base model and 2 being the competing model. Based on the MZ regression results given in Table 12, we choose HAR-RV-IV-JI, HAR-RV-JI, and HAR-CJ-JI as the baseline specifications for daily, weekly, and monthly volatility models, respectively. Hence, models providing the highest R^2 values are chosen as the baseline specifications for different forecast horizons. Besides, we consider the second-best model as the competing models. The in-sample estimation period spans from January 1, 2000 to June 30, 2018, whereas the out-of-sample period ranges from July 1, 2018 to June 30, 2020. ***Statistical significance at 1% level.

encompassing test results shown in Table 13. Note that for the encompassing tests, we choose HAR-RV-IV-JI, HAR-RV-JI, and HAR-CJ-JI as the baseline specifications for daily, weekly, and monthly volatility models, respectively. These models are selected based on the MZ regression results given in Table 12. In other words, models providing the highest R^2 values are chosen as the baseline specifications for different forecast horizons. We consider the second-best model as the competing models. Now, the results of Table 13 reveal that all the φ_1 coefficients (parameters of the baseline process) are positive and statistically significant at 1% level. Given that the parameters for the competing models appear to be insignificant, the baseline models encompass their counterparts, confirming the earlier results.

4.8 | Additional tests

So far, we have investigated whether and to what extent the JI for the VIX index improve the predictive performance of the HAR models. It is noteworthy that in addition to JI, the information on time-varying JS is also important given that it provides a good characterization of the VIX index. More specifically, the conditional JS distribution impacts the first and second-order conditional moments of the VIX index. Besides, the flexible higher-order conditional moments (i.e., skewness and kurtosis) of the VIX index can also be generated using the properties of JS distribution (Wang et al., 2022; Yin et al., 2021). Therefore, JSs for the VIX index, defined in Equation (4), could provide additional information beyond what is contained in the JI. In other words, time-varying JSs along with the JIs could play a pivotal role in forecasting the RV of S&P 500 returns. To this end, we consider two more extensions for the HAR process as given below:

$$\text{HAR-RV-JI-JS}: RV_{t,t+h} = \tau_0 + \tau_d RV_t + \tau_w RV_{t-5,t} + \tau_m RV_{t-22,t} + \tau_\lambda \lambda_t + \tau_{JS} U_{lt} + \varepsilon_t, \tag{43}$$

$$\text{HAR-CJ-JI-JS}: RV_{t,t+h} = \tau_0 + \tau\, C_t + \tau_{cw} C_{t-5,t} + \tau_{cm} C_{t-22,t} + \tau_{jd} J_t + \tau_{jw} J_{t-5,t} + \tau_{jm} J_{t-22,t} + \tau_\lambda \lambda_t + \tau_{JS} U_{lt} + \varepsilon_t. \tag{44}$$

Here, U_{lt} referring to the JS is defined in Equation (4).

Now, note that the subsample analysis, discussed in Section 4.6, shows that our proposed approach performs better than the existing models during the COVID-19 pandemic crises. However, it is stimulating to examine if such results hold for other stress or high volatility periods. To do so, we extend our subsample analysis for investigating the performance of the modified HAR model during the 2008 global financial crisis periods. In particular, our in-sample

TABLE 14 Results of MZ regression models during the 2008 crisis periods

	R^2 (%)		
Models	**Daily**	**Weekly**	**Monthly**
HAR-RV	53.4	37.1	12.9
HAR-RV-JI	58.1	39.4	14.3
HAR-RV-IV	57.2	38.0	13.5
HAR-RV-JI-JS	61.4	40.9	15.6
HAR-CJ	54.0	37.9	13.3
HAR-CJ-JI	59.1	39.8	14.9
HAR-CJ-IV	58.0	38.5	14.0
HAR-CJ-JI-JS	62.5	41.8	16.1

Note: The R^2 (%) values of the subsample analysis using the Mincer and Zarnowitz (MZ) regression model. This MZ process assumes the following form: $RV_t = \varphi_0 + \varphi_1 R\hat{V}_t + \epsilon_t$, where RV_t and $R\hat{V}_t$ refer to the true volatility and volatility forecast for day t, respectively. The sum of 5-min intra-day squared returns is used as a proxy for the true volatility. Our in-sample estimation period spans from January 1, 2000 to December 31, 2007, whereas the out-of-sample period ranges from January 1, 2008 to June 30, 2009. These results are obtained for short-term (daily), medium-term (weekly), and long-term (monthly) volatility components.

estimation period now spans from January 1, 2000 to December 31, 2007, whereas the out-of-sample period ranges from January 1, 2008 to June 30, 2009.[7]

Table 14 shows the R^2 statistics for the MZ regressions. Our objective is to compare the HAR-RV-JI-JS and HAR-CJ-JI-JS models with earlier specifications. It is now evident that the HAR-CJ-JI-JS model outperforms all other specifications while forecasting the RV during the 2008 financial crisis periods and these findings hold for daily, weekly, and monthly volatility models. Notably, this result is consistent with our prior investigation discussed in Section 4.6. It can be thus concluded that during the crisis periods including both COVID-19 pandemic and global recessions, capturing the extreme movements (e.g., jumps) in VIX plays a major role in understanding the volatility dynamics of the S&P 500 index. It is worth mentioning that the extended HAR approaches incorporating both JIs and JSs (i.e., HAR-RV-JI-JS and HAR-CJ-JI-JS models) have better predictive contents compared to HAR-RV-JI or HAR-CJ-JI models implying that JSs could provide additional information beyond what is contained in the JIs. These results also confirm the importance of considering jumps in VIX while predicting the RV.[8]

4.9 | Forecasting value-at-risk (VaR)

In this section, we aim to find the best model for predicting the 1-day ahead VaR and to serve this purpose, all the approaches are tested with a VaR forecasted under the quantile level ψ. In particular, we employ the likelihood ratio (LR) test of Kupiec (1995) and to do so, we first define the following hit sequence:

$$Hit_t = \begin{cases} 1, & if\ r_t < VaR_t \\ 0, & if\ r_t \geq VaR_t \end{cases} \tag{45}$$

where r_t denotes the return on day t and VaR_t is given by

$$VaR_t = Z_\psi \sqrt{g_t} \tag{46}$$

[7]We define the period from January 1, 2008 to June 30, 2009 as the global financial crisis period following the NBER (National Bureau of Economic Research) guidance.

[8]We provide the in-sample estimates for the HAR-RV-JI-JS and HAR-CJ-JI-JS specifications in Appendix A (see Table A2). The results suggest that both HAR-RV-JI-JS and HAR-CJ-JI-JS models outperform the earlier specifications by documenting higher R^2 statistics.

TABLE 15 VaR failure rate results

Models ↓	LQ = 10%	LQ = 5%	LQ = 1%	RQ = 10%	RQ = 5%	RQ = 1%
HAR-RV	0.509	0.299	0.167	0.662	0.543	0.801
HAR-RV-IV	0.628	0.451	0.341	0.699	0.561	0.829
HAR-RV-JI	0.661	0.499	0.400	0.778	0.849	0.887
HAR-RV-IV-JI	0.688	0.482	0.389	0.725	0.801	0.861
HAR-CJ	0.442	0.302	0.221	0.711	0.783	0.823
HAR-CJ-IV	0.598	0.442	0.366	0.681	0.592	0.917
HAR-CJ-JI	0.606	0.439	0.371	0.728	0.685	0.946
HAR-CJ-IV-JI	0.619	0.430	0.373	0.703	0.664	0.924
GJR-GARCH-IV	0.458	0.291	0.091	0.679	0.714	0.865
GJR-GARCH-JI	0.498	0.212	0.112	0.699	0.776	0.889
GJR-GARCH-IV-JI	0.498	0.249	0.149	0.682	0.918	0.875
HAR-RV-JI-JS	0.722	0.529	0.433	0.798	0.937	0.971
HAR-CJ-JI-JS	**0.731**	**0.557**	**0.459**	**0.811**	**0.961**	**0.972**

Note: This Table reports the p values of the Kupiec (1995) failure rate test. The larger the p value, the higher the accuracy of the VaR predicted by the model. We use the rolling window method to forecast the 1-day ahead VaR. The out-of-sample forecasting horizon is the COVID-19 pandemic period. Therefore, we use a rolling window from January 1, 2016 to December 31, 2019. We provide the empirical failure rates for left quantile (LQ) and right quantile (RQ) at 10%, 5%, and 1%. Bold refers to the best model for VaR forecasting under a specific quantile level.

with Z_ψ being the quantile at $100 \times \psi\%$ of the standardized probability distribution and g_t indicating the volatility that can be estimated by different competing models (see Giot & Laurent, 2003, 2004; Giot, 2005).

Now, let N be the number of VaR violations and T refer to the total number of observations. We then wish to test the null hypothesis $H_0 : f = \psi$, with f being the failure rate. In line with Kupiec (1995), we construct the LR test statistic as follows:

$$LR = -2\ln\left\{(1-\psi)^N \psi^{T-N} \Big/ \left(1-\frac{N}{T}\right)^{T-N}\left(\frac{N}{T}\right)^N\right\} \sim \chi^2(1). \tag{47}$$

Table 15 reports the p values of the above likelihood ratio test. We use the rolling window method to forecast the 1-day ahead VaR. Note that the out-of-sample forecasting horizon is the COVID-19 pandemic period and we use a rolling window from January 1, 2016 to December 31, 2019 so that we have a moving window of at least 1000 observations. We provide the empirical failure rates for left quantile (LQ) and right quantile (RQ) at 10%, 5% and 1%. We have compared 13 different models in the VaR analysis: HAR-RV, HAR-RV-IV, HAR-RV-JI, HAR-RV-IV-JI, HAR-CJ, HAR-CJ-IV, HAR-CJ-JI, HAR-CJ-V-JI, GJR-GARCH-IV, GJR-GARCH-JI, GJR-GARCH-IV-JI, HAR-RV-JI-JS and HAR-CJ-JI-JS. The findings show that our proposed model HAR-CJ-JI-JS appear to be best approach for VaR forecasting in each of the cases considered. There is only one case (see the results at RQ = 1%) where both HAR-RV-JI-JS and HAR-CJ-JI-JS produce similar predictions. Overall, we find that both JSs and JIs for VIX have a significant role to play in predicting the 1-day ahead VaR.

4.10 | Discussion

Overall, our findings indicate that the extended HAR models which incorporate JIs and JS seem to perform a reasonable job. In fact, the full period analysis confirms that the HAR-RV-JI, HAR-CJ-JI, HAR-RV-JI-JS, and HAR-RV-JI-JS models are found to surpass all the existing HAR-RV specifications while predicting future RV of the S&P 500 index. However, it is also observed that the performance of these extended HAR-RV models tends to vary under diverse market conditions. In particular, we find evidence that during periods of market stress, the extended models generate more accurate out-of-sample forecasts than do the existing HAR approaches. Amid the low volatility periods, on the other hand, our results are somewhat inconclusive, albeit the HAR-type models incorporating jumps in the VIX index still provide more accurate out-of-sample forecasts than do the existing approaches.

These results are not unexpected given that jumps in VIX allow volatility to rapidly increase which provides some signals of market crash. For example, in the market stress of March 2020, VIX has jumped roughly 33% and consequently, the S&P 500 index has seen the biggest plunge since the crash on Black Monday in 1987. Similar jumps in the VIX index are observed during the 2008 financial downturn as well. Eraker et al. (2003) also show that when compared to jumps in returns, jumps in VIX plays a more prominent role in generating market stress. These facts simply indicate that jumps in VIX are a primary component that generates chaos in financial markets. Therefore, the proposed HAR-RV models which include the JI parameter offer further improvements in the accuracy of the volatility forecasts.

It is also noteworthy that unlike the JS, the expected number of jumps (i.e., JI) is largely clustered, exhibiting greater persistence and predictability (Aït-Sahalia & Hurd, 2015). This further confirms the fact that the intensity of jumps in VIX may have superior predictive contents for stock price volatility during the stress periods (Ma et al., 2014). While the magnitude (or size) of volatility jumps is extensively studied in previous studies, JI has received relatively less attention. Economically, a higher likelihood of extreme movements in VIX implies less stability in the market and hence, the stock price volatility in the next period tends to be larger. Empirically, prior literature (see Clements & Liao, 2017) demonstrates that due to the persistence of JI in regard to VIX, the RV of S&P 500 returns becomes more predictable amid the turmoil periods.

In sum, jumps in VIX play an important role as they represent a significant source of nondiversifiable risk. Since policymakers and investors must make decisions in real-time during periods of jump-inducing chaotic conditions in financial markets, it is economically important to adopt appropriate volatility models which can capture the impact of jumps in the VIX index. Hence, our proposed models could be of interest to those who attempt to predict volatility during the periods of market stress.

5 | CONCLUSIONS

Prior literature documents that jumps occurring in the VIX index lead to rapid changes in the level of volatility, which in turn raise the probability of extreme movements in the US stock market. This paper empirically shows the importance of using the information content of volatility jumps in forecasting equity return volatility. In particular, our analysis reveals that jumps in VIX have a positive impact on the RV of S&P 500 index and that the evidence of such jumps provides incremental forecasting information for the US stock market. We further show that the proposed HAR-RV approach generates more accurate out-of-sample volatility forecasts than do the existing HAR-RV models. Importantly, these results hold for short-, medium-, and long-term volatility components. Our findings thus demonstrate that as the volatility jumps tend to increase the uncertainty in stock markets, they could have significant predictive content for equity return volatility.

Overall, the results reveal that the information on time-varying volatility jumps could play a major role in analyzing the risk of investor portfolio. Thus, our findings seem important for volatility modeling and portfolio constructions in the presence of time-varying jumps. Given that capturing jumps in VIX plays a significant role in risk management, option pricing and portfolio optimization, this paper has important implications to investors holding assets in the US equity market. Policymakers may also utilize these results to develop appropriate hedging strategies amid the period of high uncertainties.

Moreover, this empirical research has important implications to investors given that our findings can be exploited to understand the economics of the market further. As we find that time-varying jumps in implied volatilities provide additional information beyond what is contained in VIX, the informational content of JIs and JSs could be useful for making proper asset allocation decisions. It is also worth noting that investors tend to react differently to positive and negative shocks, which may arrive independently with different rates and sizes. Therefore, the information on volatility jumps, which capture such shocks, could be useful for improving the forecast accuracy of the US equity market volatility under diverse market conditions. Since jumps in VIX lead to a sudden increment in the levels of volatility, it is particularly important for risk managers to recognize the impact of such jumps on the dependence structure between VIX and stock returns. The recent downturns in financial markets have also emphasized the significance of volatility jumps for understanding the various forms of risk inherent in stock returns and their implications for portfolio optimization. To this end, our study could be important for investors as the findings show that the information on volatility jumps is useful for the improvement in the accuracy of volatility forecasts during the stress periods.

ACKNOWLEDGMENT

The authors do not receive any funding for this study work.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available at Oxford-Man Institute's Realized Library version 0.3 (https://realized.oxford-man.ox.ac.uk) and in the website of Chicago Board Options Exchange (https://ww2.cboe.com).

ORCID

Anupam Dutta http://orcid.org/0000-0003-4971-3258

REFERENCES

Ahmad, W., Sadorsky, P., & Sharma, A. (2018). Optimal hedge ratios for clean energy equities. *Economic Modelling*, *72*, 278–295.

Aït-Sahalia, Y., & Hurd, T. R. (2015). Portfolio choice in markets with contagion. *Journal of Financial Econometrics*, *14*(1), 1–28.

Akgiray, V. (1989). Conditional heteroscedasticity in time series of stock returns: Evidence and forecasts. *The Journal of Business*, *62*, 55–80.

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The Review of Economics and Statistics*, *89*(4), 701–720.

Andersen, T. G., Bollerslev, T., & Huang, X. (2011). A reduced form framework for modeling volatility of speculative prices based on realized variation measures. *Journal of Econometrics*, *160*(1), 176–189.

Baldeaux, J., & Badran, A. (2014). Consistent modelling of VIX and equity derivatives using a 3/2 plus jumps model. *Applied Mathematical Finance*, *21*(4), 299–312.

Bandi, F. M., & Reno, R. (2016). Price and volatility co-jumps. *Journal of Financial Economics*, *119*(1), 107–146.

Barndorff-Nielsen, O. E., & Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, *2*(1), 1–37.

Barndorff-Nielsen, O. E., & Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, *4*(1), 1–30.

Basher, S. A., Haug, A. A., & Sadorsky, P. (2016). The impact of oil shocks on exchange rates: A Markov-switching approach. *Energy Economics*, *54*, 11–23.

Becker, R., Clements, A. E., & White, S. I. (2006). On the informational efficiency of S&P500 implied volatility. *The North American Journal of Economics and Finance*, *17*(2), 139–153.

Bera, A. K., & Higgins, M. L. (1993). ARCH models: Properties, estimation and testing. *Journal of Economic Surveys*, *7*(4), 305–366.

Bollerslev, T. (1987). A conditional heteroskedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics*, *69*, 542–547.

Bollerslev, T., Chou, R. Y., & Kroner, K. F. (1992). ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, *52*(1–2), 5–59.

Bollerslev, T., Engle, R. F., & Nelson, D. B. (1994). ARCH models. *Handbook of Econometrics*, *4*, 2959–3038.

Brailsford, T. J., & Faff, R. W. (1996). An evaluation of volatility forecasting techniques. *Journal of Banking & Finance*, *20*(3), 419–438.

Busch, T., Christensen, B. J., & Nielsen, M. Ø. (2011). The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets. *Journal of Econometrics*, *160*(1), 48–57.

Canina, L., & Figlewski, S. (1993). The informational content of implied volatility. *The Review of Financial Studies*, *6*(3), 659–681.

Carr, P., & Wu, L. (2006). A tale of two indices. *The Journal of Derivatives*, *13*(3), 13–29.

Çelik, S., & Ergin, H. (2014). Volatility forecasting using high frequency data: Evidence from stock markets. *Economic Modelling*, *36*, 176–190.

Chan, W. H., & Maheu, J. M. (2002). Conditional jump dynamics in stock market returns. *Journal of Business & Economic Statistics*, *20*(3), 377–389.

Chiang, S.-M., Chen, C.-D., & Huang, C.-M. (2019). Analyzing the impacts of foreign exchange and oil price on biofuel commodity futures. *Journal of International Money and Finance*, *96*, 37–48.

Chong, Y. Y., & Hendry, D. F. (1986). Econometric evaluation of linear macro-economic models. *The Review of Economic Studies*, *53*(4), 671–690.

Christensen, B. J., & Prabhala, N. R. (1998). The relation between implied and realized volatility. *Journal of Financial Economics*, *50*(2), 125–150.

Clements, A., & Liao, Y. (2017). Forecasting the variance of stock index returns using jumps and cojumps. *International Journal of Forecasting*, *33*(3), 729–742.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, *7*(2), 174–196.

Corsi, F., Pirino, D., & Reno, R. (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics*, *159*(2), 276–288.

Day, T. E., & Lewis, C. M. (1992). Stock market volatility and the information content of stock index options. *Journal of Econometrics*, *52*(1–2), 267–287.

Dennis, P., Mayhew, S., & Stivers, C. (2006). Stock returns, implied volatility innovations, and the asymmetric volatility phenomenon. *Journal of Financial and Quantitative Analysis*, *41*(2), 381–406.

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

Dimson, E., & Marsh, P. (1990). Volatility forecasting without data-snooping. *Journal of Banking & Finance*, *14*(2–3), 399–421.

Dutta, A., Bouri, E., & Roubaud, D. (2021). Modelling the volatility of crude oil returns: jumps and volatility forecasts. *International Journal of Finance & Economics*, *26*(1), 889–897.

Eraker, B., Johannes, M., & Polson, N. (2003). The impact of jumps in volatility and returns. *The Journal of Finance*, *58*(3), 1269–1300.

Fernandes, M., Medeiros, M. C., & Scharth, M. (2014). Modeling and predicting the CBOE market volatility index. *Journal of Banking & Finance*, *40*, 1–10.

Figlewski, S. (1997). Forecasting volatility. *Financial Markets, Institutions & Instruments*, *6*(1), 1–88.

Forsberg, L., & Ghysels, E. (2007). Why do absolute returns predict volatility so well? *Journal of Financial Econometrics*, *5*(1), 31–67.

Franses, P. H., & Van Dijk, D. (1996). Forecasting stock market volatility using (non-linear) Garch models. *Journal of forecasting*, *15*(3), 229–235.

Giot, P. (2005). Relationships between implied volatility indexes and stock index returns. *The Journal of Portfolio Management*, *31*(3), 92–100.

Giot, P., & Laurent, S. (2003). Value-at-risk for long and short trading positions. *Journal of Applied Econometrics*, *18*(6), 641–663.

Giot, P., & Laurent, S. (2004). Modelling daily value-at-risk using realized volatility and ARCH type models. *Journal of Empirical Finance*, *11*(3), 379–398.

Giot, P., & Laurent, S. (2007). The information content of implied volatility in light of the jump/continuous decomposition of realized volatility. *The Journal of Futures Markets*, *27*(4), 337–359.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, *48*(5), 1779–1801.

Gronwald, M. (2019). Is Bitcoin a commodity? On price jumps, demand shocks, and certainty of supply. *Journal of International Money and Finance*, *97*, 86–92.

Kambouroudis, D. S., & McMillan, D. G. (2016). Does VIX or volume improve GARCH volatility forecasts? *Applied Economics*, *48*(13), 1210–1228.

Kat, H. M., & Heynen, R. C. (1994). Volatility prediction: A comparison of the stochastic volatility, GARCH (1, 1) and EGARCH (1, 1) models. *The Journal of Derivatives*, *2*(2), 50–65.

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives Winter*, *3*(2), 73–84.

Ma, F., Wei, Y., Huang, D., & Chen, Y. (2014). Which is the better forecasting model? A comparison between HAR-RV and multifractality volatility. *Physica A: Statistical Mechanics and Its Applications*, *405*, 171–180.

Mincer, J. A., & Zarnowitz, V. (1969). The evaluation of economic forecasts, Economic forecasts and expectations: Analysis of forecasting behavior and performance (pp. 3–46). NBER.

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

Nelson, D. B. (1992). Filtering and forecasting with misspecified ARCH models I: Getting the right variance with the wrong model. *Journal of Econometrics*, *52*(1–2), 61–90.

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

Pati, P. C., Barai, P., & Rajib, P. (2018). Forecasting stock market volatility and information content of implied volatility index. *Applied Economics*, *50*(23), 2552–2568.

Thomson Reuters DataStream database. Oxford-Man Institute's Realized Library version 0.3.

Wang, T., Cheng, S., Yin, F., & Yu, M. (2022). Directly pricing VIX futures: The role of dynamic volatility and jump intensity. *Applied Economics*, *54*, 3678–3694.

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

Yu, W. W., Lui, E. C. K., & Wang, J. W. (2010). The predictive power of the implied volatility of options traded OTC and on exchanges. *Journal of Banking & Finance*, *34*(1), 1–11.

Zhou, C., Wu, C., & Wang, Y. (2019). Dynamic portfolio allocation with time-varying jump risk. *Journal of Empirical Finance*, *50*, 113–124.

How to cite this article: Dutta, A., & Das, D. (2022). Forecasting realized volatility: New evidence from time-varying jumps in VIX. *Journal of Futures Markets*, 1–25. https://doi.org/10.1002/fut.22372

APPENDIX A

TABLE A1 Summary statistics of jump intensities for the VIX index

Mean	SD	Maximum	Minimum
0.1424	0.1076	1.1172	0.0043

Note: The summary statistics of jump intensities (i.e., λ_t) for the VIX index.

TABLE A2 Impact of jump intensities and jump sizes on RV

Models	HAR-RV-JI-JS Daily	HAR-RV-JI-JS Weekly	HAR-RV-JI-JS Monthly	HAR-CJ-JI-JS Daily	HAR-CJ-JI-JS Weekly	HAR-CJ-JI-JS Monthly
τ_0	2.8×10^{-5}*** (1.9×10^{-6})	8.1×10^{-5}** (3.9×10^{-5})	-4.0×10^{-5}*** (1.8×10^{-6})	2.2×10^{-5}*** (4.1×10^{-6})	1.6×10^{-5}*** (5.9×10^{-6})	-4.3×10^{-5}*** (2.4×10^{-6})
τ_d	0.2943*** (0.0261)	0.1892*** (0.0248)	0.1273*** (0.0235)			
τ_w	0.4100*** (0.0537)	0.2843*** (0.0311)	0.2876*** (0.0411)			
τ_m	0.3211*** (0.0878)	0.3389*** (0.0659)	0.2994*** (0.0338)			
τ_{cd}				0.3651*** (0.0256)	0.2573*** (0.0476)	0.1769*** (0.0335)
τ_{cw}				0.5983*** (0.0113)	0.4801*** (0.0311)	0.2791*** (0.0738)
τ_{cm}				−0.1309*** (0.0336)	−0.1178*** (0.0389)	−0.0988*** (0.0302)
τ_{jd}				−0.1457*** (0.0288)	−0.1424*** (0.0373)	−0.0879** (0.0410)
τ_{jw}				−0.6999*** (0.0877)	−0.7152*** (0.0971)	−0.2611*** (0.0644)
τ_{jm}				1.8119*** (0.2259)	1.8834*** (0.1829)	1.4561*** (0.1864)
τ_λ	2.0×10^{-4}*** (2.3×10^{-5})	2.9×10^{-4}*** (2.4×10^{-5})	2.6×10^{-4}*** (5.1×10^{-5})	1.9×10^{-4}*** (3.3×10^{-5})	3.0×10^{-4}*** (2.9×10^{-5})	2.8×10^{-4}*** (4.2×10^{-5})
τ_{JS}	1.0×10^{-5}*** (1.8×10^{-6})	1.1×10^{-5}*** (2.0×10^{-6})	1.6×10^{-5}*** (4.1×10^{-6})	1.1×10^{-5}*** (2.0×10^{-6})	1.4×10^{-5}*** (2.1×10^{-6})	1.5×10^{-5}*** (3.6×10^{-6})
R^2 (%)	58.2	63.0	57.9	59.3	66.1	58.8
HET test	0.28	0.51	0.48	0.23	0.61	0.79

Note: The in-sample estimates for the HAR models defined in Equations (43) and (44). These models incorporate both jump intensities and jump sizes for the VIX index. The in-sample estimation period spans from January 1, 2000 to June 30, 2018, yielding a total of 4654 daily observations. Values in parentheses indicate the standard errors. We also provide the estimates of R^2 statistics. The values shown in the last row refer to the p values for the F test of no heteroscedasticity. ***, **, and * imply that the null hypothesis is rejected at 1%, 5%, and 10% significance levels, respectively.