VIX Forecasting via HAR Volatility Forecasting Framework

SICHENG WANG

MASTER OF FINANCIAL ENGINEERING

CORNELL UNIVERSITY

sw863@cornell.edu

December 2, 2024

Abstract

This study utilizes daily realized volatility measures to improve forecast accuracy for stock market implied volatility, using the daily VIX Index as proxy, across forecast horizons of 1, 5, and 22 days. The heterogeneous autoregressvie (HAR) model was used as the forecasting framework in this study. The HAR model is often estimated via ordinary least squares (OLS). However, given the limitations of OLS in dealing with outliers, conditional heteroskedasticity, and non-Gaussianity, remedies such as an additional $GARCH$ Error term or estimation via Weighted Least Square (WLS) were investigated in this study. The aim of this study is to use the modified HAR model to model and capture the most recent single day VIX spike on August 5th, 2024, and to see how much improvement can the modified model achieve compare to the original HAR model.

1 Introduction

Due to the shortcomings of the $GARCH$ model in capturing the key empirical features of financial returns (long memory, fat tails, and self-similarity). Corsi (2009) propose an additive cascade model [1] incorporating different volatility components with different time horizons. According to Corsi (2009), market volatility can be divided into three main components: the high frequency and day trader with daily or higher trading frequency; the medium frequency investors who typically rebalance weekly, and the low frequency agents with investment time-horizon of one or more months.

The HAR model can often be estimated via ordinary least squares regression (OLS). However, given the characteristics of Volatility data (VIX index is used as a proxy in this paper) such as outliers, conditional heteroskedasticity, non-Gaussianity, and taken into considerations of the limitations of the

[1] An additive cascade model in econometrics refers to a model where multiple layers or stages of effects are added together to explain the structure of a process or variable, often across different levels of a system or over time. This concept is commonly applied in modeling time series, spatial data, or hierarchical structures.

OLS model such as highly sensitive to outliers, poor estimation with conditional heteroskedasticity or non-Gaussianity, using OLS to estimate and forecast Market Volatility (VIX) is not an optimal choice, leaving room for improvements (Clements and Preve, 2021).

To address the short comings of OLS estimator, Huang et al. (2016) proposed an $HAR-GARCH$ model to use the $GARCH$ to account for the Latent volatility in the HAR model, providing an improvement by accounts for conditional heteroskedasticity of the volatility data while preserving the benefit of long-memory of the original HAR model. On the other hand, Clements and Preve (2021) provided an less complicated remedy to the HAR model by implementing it via the Weighted Least Square estimator (WLS). Clements and Preve (2021) achieved considerable reductions in common statistical and economic loss measures compare to the estimation via OLS. To investigate the validity of past studies and to explore better method in forecasting Volatility data, $HAR-GARCH$ model and Weight least Square HAR model are investigated in this paper to model the most recent market spikes on August 5th, 2024.

2 Data

The key data in this study consists of daily observations of the VIX volatility index computed by the Chicago Board Options Exchange. First introduced in 1993, the VIX is calculated via the bid/ask quotes of options on the S&P 500 index. The VIX index is not only viewed as the market expected volatility but also an indicator of investor sentiment and risk aversion. Assume investors concerns about the market outlook, the demand for S&P 500 put option will rise, resulting an increase in both Implied Volatility and the value of the VIX (Ahoniemi, 2008).

The computation method for VIX was modified on September 22, 2003 where the original calculation based on S&P 100 options was changed to S&P 500 options, thus making VIX a more robust proxy for market volatility. Additionally, calculation of VIX is independent of the option pricing model (ex. Black Scholes) where it is computed directly using the two nearest expiring S&P 500 options (monthly options) to cover a 30-calendar-dat-period. More specifically, the VIX index is calculated via the prices of at-money and out-of-money calls and puts. The equation can be expressed as:

$$\sigma^2 = \frac{2}{T}\sum_i \frac{\Delta K_i}{K_i^2} e^{RT} Q(K_i) - \frac{1}{T}[\frac{F}{K_0} - 1]^2$$

where $VIX = \sigma \times 100$, T represent the time to expiration, F is the Forward index level derived from index option prices, K_i is the strike price of the i_{th} out-of-money option (call if $K_t > F$ and put if

$K_t < F$), ΔK_i is the interval between strike prices ($(K_{i+1} - K_{i-1})/2$), K_0 is the first strike below F, R is the risk-free interest rate to expiration, and $Q(K_i)$ is the midpoint of the bid-ask spread for option with strike K_i (Ahoniemi, 2008).

3 Model

3.1 Modeling the VIX

As we can observe from the Figure 1 below, we can see the VIX index was somewhat stable in the early 1990s until 1996. Spikes in the VIX before 2000s coincide with the Asian financial crisis of late 1997 and the Russian Default LTCM crisis of late 1998. The two largest spikes in the VIX history was the 2008 Financial Crisis and the Covid-19 Pandemic, reflecting when market uncertainty reached the peak and investor confidence to the lowest. In this paper, the goal is to model the most recent spikes in the VIX index, which was the sell-off on August 5th, 2024 due to a combination of weak US economic data, an interest rate hike in Japan and signs of overvaluation in tech stocks (Ahoniemi, 2008).

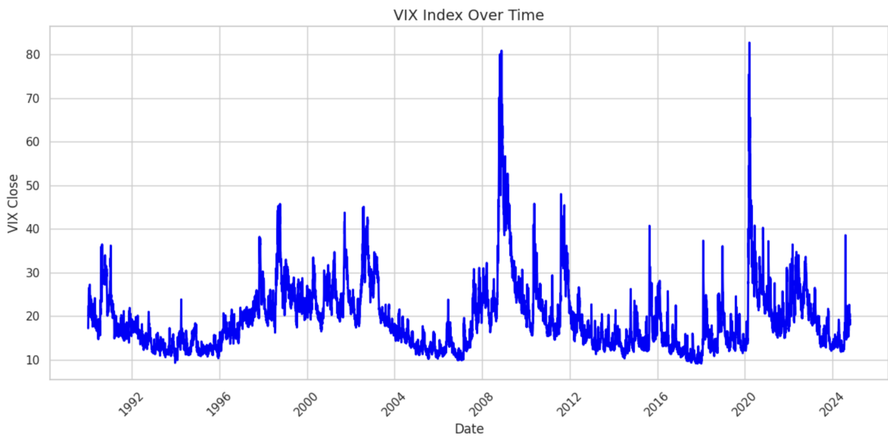


Figure 1: VIX index 1.1.1990 - 24.10.2024

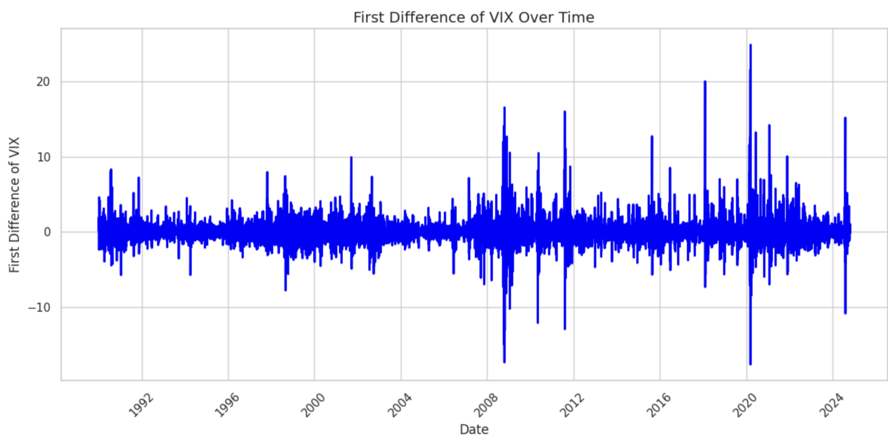


Figure 2: VIX First Difference 1.1.1990 - 24.10.2024

From Figure 2 above, we can see the VIX first differences which serve as a visual inspection of heteroskedasticity in the VIX data.

Natural logarithms were applied to VIX data to avoid negative forecasts of volatility. The log-VIX display high level of autocorrelation, as shown in Figure 3. The ACF of log VIX first difference shows little autocorrelation.

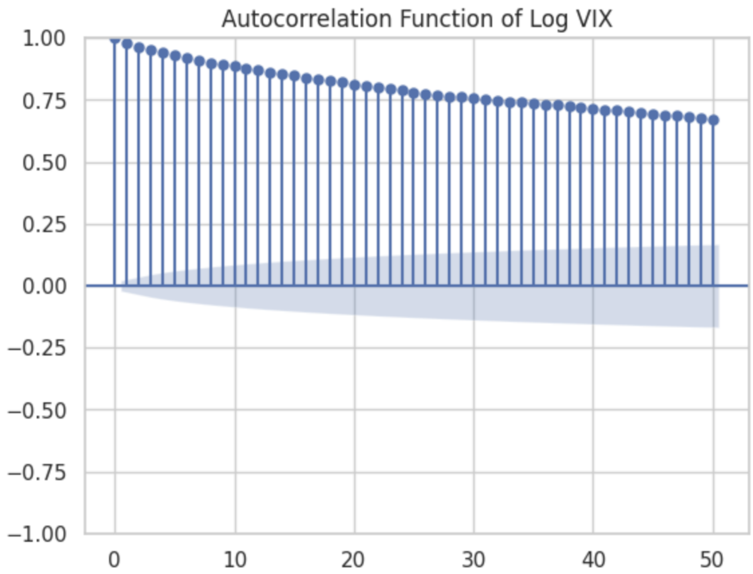


Figure 3: Log VIX ACF

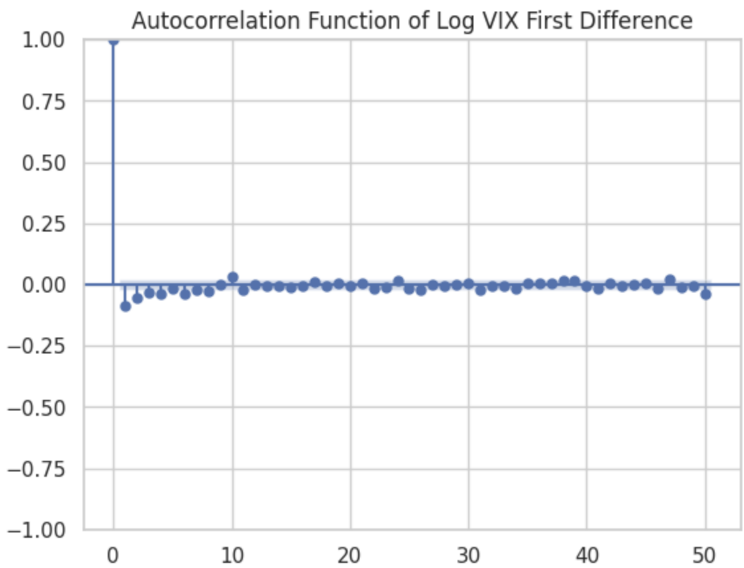


Figure 4: Log VIX First Difference ACF

3.2 The Autoregressive Model

To construct a simple benchmark model, the autoregressive model with 1 day lag ($AR1$) is implemented. Compared to other naive models available, $AR1$ model is chosen as benchmark due to its close resemblance to the HAR framework, which incorporates extra lags to account for volatility relative to the $AR1$ model (Degiannakis and Kafousaki, 2023). The $AR1$ model can be expressed as:

$$log(\hat{VIX_t}) = \beta_1 + \beta_2 log(VIX_{t-1}) + \epsilon_t$$

3.3 The HAR Model

This additive cascade volatility forecasting model proposed by Corsi (2009) is called the Heterogeneous AutoRegressive model (HAR), in which an asset's t realized daily variance RV_t (model via $log(VIX)_t$) is modeled as an AR(22) process, given formula:

$$log(\hat{VIX})_t = \beta_0 + \beta_d VIX_{t-1} + \beta_w(5^{-1}\sum_{k=1}^{5} log(VIX_{t-k})) + \beta_m(22^{-1}\sum_{k=1}^{22} log(VIX_{t-k})) + \epsilon_t$$

, where:

- $log(\hat{VIX})_t$ is the asset's forecast log VIX at time t
- $log(VIX)_{t-1}$ is the daily realized variance represented by lagged 1 daily log VIX at time t
- $(5^{-1}\sum_{k=1}^{5} log(VIX_{t-k}))$ is the asset's weekly realized variance represented by the average of lagged 5 dalily log VIX at time t
- $(22^{-1}\sum_{k=1}^{22} log(VIX_{t-k}))$ is the asset's monthly realized variance represented by the lagged 22 daily log VIX at time t

3.4 Realized HAR GARCH Model

The Realized $GARCH$ model builds upon the traditional $GARCH$ by taken into account realized measures (e.g., realized volatility) to improve the modeling of latent volatility (Huang et al., 2016). The model consists of three key equations: Return Equations, $GARCH$ Equation, and Measurement Equation.

1. Return Equation

$$r_t = \mu + \sqrt{h_t} z_t, z_t \sim N(0,1)$$

   - r_t: Asset returns at time t
   - h_t: Conditional variance
   - z_t: Standardized residuals

2. GARCH Equation

$$log(h_t) = w + \beta \times log(h_{t-1} + \rho \times log(x_{t-1}))$$

- h_t: Latent volatility
- x_{t-1}: Realized measure (e.g., realized volatility or kernel)

3. Measurement Equation

$$log(x_t) = \xi + \phi \times log(h_t) + \tau(z_t) + \mu_t, \mu_t \sim N(0, \sigma_u^2)$$

where it links observed realized measures to latent volatility.

4. Leverage Function

$$\tau(z_t) = \tau_1 z_t + \tau_2(z_t^2 - 1)$$

Captures asymmetric dependence between returns and volatility.

3.4.1 Realized HAR-GARCH Model

According to Huang et al. (2016) The Realized $HAR - GARCH$ model extends the Realized GARCH model by incorporating the Heterogeneous Autoregressive (HAR) framework. It account for long memory and persistence in volatility through multiple lags of realized measures (daily, weekly, and monthly components).

(a) Latent Volatility Equation

$$log(h_t) = \omega + \beta log(h_{t-1}) + \gamma_d log(x_{t-1}) + \frac{\gamma_w}{4} \sum_{i=2}^{5} log(x_{t-i}) + \frac{\gamma_m}{17} \sum_{i=6}^{22} log(x_{t-i})$$

where $\gamma_d, \gamma_w, \gamma_m$ are coefficients for daily, weekly, and monthly realized volatility components (Modeled via VIX in this study).

(b) Reparameterized Form

$$log(h_t) = \omega + \beta log(h_{t-1}) + \gamma_d^* log(x_{t-1}) + \frac{\gamma_w^*}{5} \sum_{i=1}^{5} log(x_{t-i}) + \frac{\gamma_m^*}{22} \sum_{i=1}^{22} log(x_{t-i})$$

Reparameterization simplifies estimation by consolidating coefficients.

(c) Measurement Equation

$$log(x_t) = \mu_x + \phi log(h_t) + \tau(z_t) + u_t$$

- $\tau(z_t)$: Leverage function
- u_t: Independent noise

3.5 HAR Model via Weighted Least Square Estimation

The Weighted Least Square (WLS) method modifies the HAR model by multiplying both sides of the equation by the square root of the weights $w_t = \frac{1}{\sigma_t^2}$, where σ_t^2 is the estimated variance of residuals from OLS. The equation of HAR via WLS is given by:

$$\sqrt{w_t}log(VIX_t) = \sqrt{w_t}\beta_0 + \sqrt{w_t}\beta_d log(VIX_{t-1}) + \sqrt{w_t}\beta_w \frac{1}{5}\sum_{i=1}^{5} log(VIX_{t-i})$$
$$+ \sqrt{w_t}\beta_m \frac{1}{22}\sum_{i=1}^{22} log(VIX_{t-i}) + \sqrt{w_t}\epsilon_t$$

4 Empirical Results

4.1 In-sample results

While forecasting the most recent one day spike using the HAR models is the focus of this paper, this section outlines various in-sample estimation results to provide some insights into the important features of the estimation. Table 1 below presents the in-sample estimation for the $AR(1)$, HAR, $HAR - GARCH$, and WLS_{HAR}. From the benchmarks $AR(1)$ model, we can see that the 1-day lag ($VIXR_D$) is a very strong predictor. Even with just one predictor, the AR(1) model was able to achieve an in-sample R^2 of 0.963. Then by adding in the Weekly and Monthly lag, we arrived at the standard HAR model. However, for the HAR model, the majority weight is still concentrated on the 1-day lag predictor, and the in-sample R^2 improved slightly to 0.964 from the benchmark model. The HAR model was then augmented with Realized $GARCH$ errors to account for latent volatility to construct the $HAR - GARCH$ model. The estimation for the $GARCH$ coefficients are all statistically significant, however, the additional $GARCH$ error term did not seem to improve the R^2 of the model compare the benchmark $AR(1)$ and original HAR model. Finally, for the Weighted Least Square estimation of HAR (WLS_{HAR}), we can see from Table 1 below that the weights of 1-day lag has been reduced significantly and weights of monthly lag increased significantly, while all the coefficient remains significant. Additionally, we can see some improvement in R^2 of WLS_{HAR} compare to benchmark $AR(1)$ and the original HAR model.

4.2 Out Sample Results

Forecasts were calculated from all models presented above. In this study, total 8763 daily observation of VIX is used, while 8600 were used in-sample training, 113 were used for out-of-sample forest, and 50 sample were used as purge in between in-sample and out-sample to prevent information leakage. The 113 out-of-sample include the largest recent one day spike in VIX on August 5th, 2024. Thus to accurately forecast the out-sample set and the spike on August 5th, 2024 was the main goal of the forecasting.

Table 1: In Sample Estimation Results

	AR 1	**HAR**	**HAR-GARCH**	**WLS-HAR**
const	0.0537	0.0354	0.0354	0.0318
	(0.00)	(0.00)	(0.00)	(0.00)
VIXR_D	0.9815	0.8871	0.8871	0.8435
	(0.00)	(0.00)	(0.00)	(0.00)
VIXR_W	-	0.0611	0.0611	0.0413
		(0.00)	(0.00)	(0.00)
VIXR_M	-	0.0396	0.0396	0.1038
		(0.00)	(0.00)	(0.00)
omega	-	-	0.00043485	-
			(0.00)	
alpha_1	-	-	0.1	-
			(0.00)	
beta_1	-	-	0.8	-
			(0.00)	
R2̂	0.963	0.964	0.928	0.972
In sample MSE	0.004413	0.004348	0.008695	0.004403
In sample QLIKE	2.185	2.155	4.382	2.179

Patton (2011) finds that the mean square error (MSE) and the empirical quasi-likelihood (Q-LIKE) are the most robust metrics for forecasting model evaluation. Thus, in this study, we deployed these two loss functions:

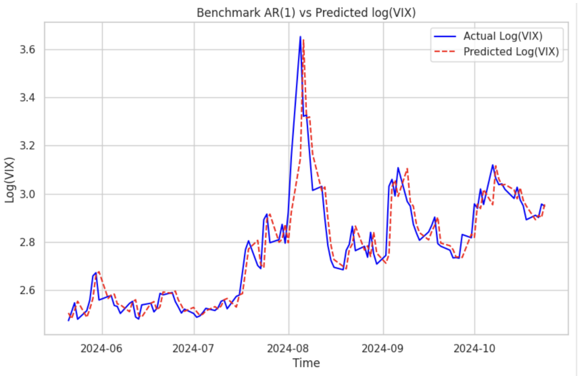
$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2$$

$$Q - LIKE = \frac{1}{n}\sum_{i=1}^{n}(log(\hat{y}_i + \frac{y_i}{\hat{y}_i}))$$

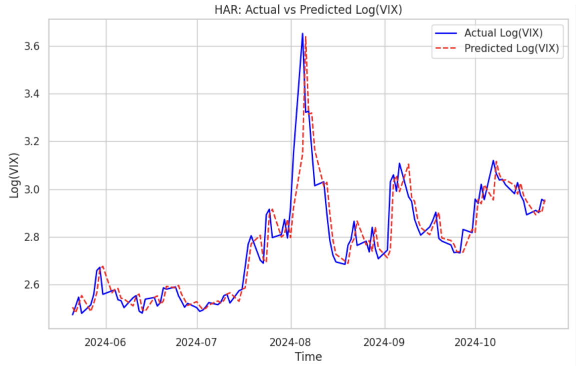
From the out-of-sample performance metrics result in table 2 we can see that, WLS_{HAR} provides some improvements from the benchmark $AR(1)$ model and the original HAR model, making Weighted Least Square a sounds choice for HAR estimation. Thus, WLS_{HAR} provides a partial remedy to the limited OLS estimation. Accoring to Clements and Preve (2021), the improvement moving from OLS to WLS for HAR estimation is like due to less over-prediction of WLS compare to OLS, with lower average under-predictions and absolute over-predictions. On the other hand, $HAR-GARCH$ significantly under-performs both the bench market $AR(1)$ model and the original HAR model, adding some questions to whether it is worth the value to add the redundant $GARCH$ error term.

Table 2: Out-of-Sample Performance Metrics

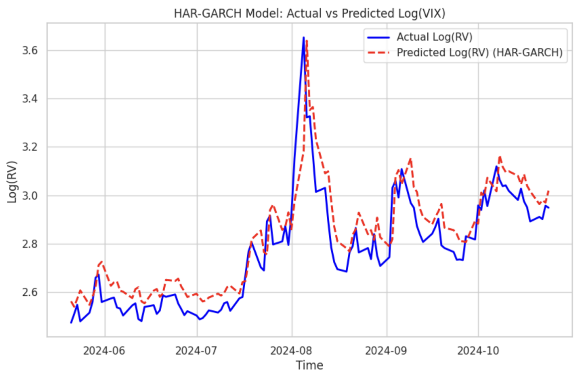
	AR 1	**HAR**	**HAR-GARCH**	**WLS-HAR**
Out-Sample MSE	0.008246	0.008166	0.011812	0.008133
Out-Sample QLIKE	0.051	0.051	0.076	0.051



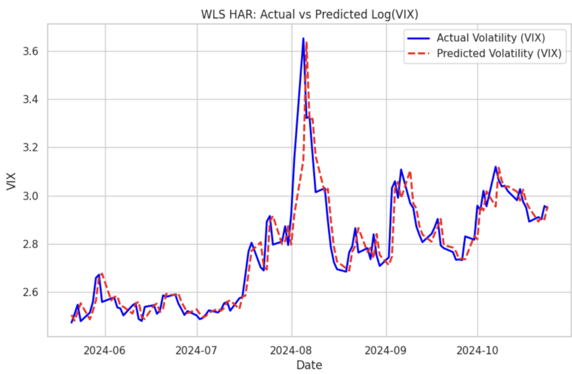
(a) AR(1)



(b) HAR



(c) HAR-GARCH



(d) WLS-HAR

Figure 5: Forecasting Results

5 Conclusion

This study sought to improve the forecasting performance of the standard HAR model for the VIX index. For the standard HAR model, the 1-day lagged $log(VIX)$, the 1-week lagged $log(VIX)$, and the 1-month lagged $log(VIX)$ all serve as statistically significant explanatory variables for $log(VIX)$. Realized $GARCH$ errors term are statistically significant in the $HAR - GARCH$ model, however, the realized $GARCH$ errors do not improve the forecast accuracy of the various models considered. The best models forecast the direction of change of the VIX is the HAR model via the Weighted Least Square estimation, leading to improvement in in-sample R^2 as well as out-of-sample MSE. The improvements using WLS over OLS can be contributed to less over-prediction of WLS compare to OLS.

A number of interesting avenues for futures research arise. a more detailed fine tune of the Realized $GARCH$ term and combine it with the HAR framework to account for latent volatility can be further investigated. Additionally, machine learning methods such as Support Vector Regression (SVR) can also be explored with the HAR framework to further improve the forecast accuracy of the HAR models.

References

K. Ahoniemi. Modeling and forecasting implied volatility - an econometric analysis of the vix index. 07 2008. doi: 10.2139/ssrn.1033812.

A. Clements and D. P. Preve. A practical guide to harnessing the har volatility model. *Journal of Banking Finance*, 133:106285, 2021. ISSN 0378-4266. doi: https://doi.org/10.1016/j.jbankfin.2021.106285. URL https://www.sciencedirect.com/science/article/pii/S0378426621002417.

F. Corsi. A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics*, 7(2):174–196, 02 2009. ISSN 1479-8409. doi: 10.1093/jjfinec/nbp001. URL https://doi.org/10.1093/jjfinec/nbp001.

S. Degiannakis and E. Kafousaki. Forecasting VIX: The illusion of forecast evaluation criteria. Working Papers 322, Bank of Greece, June 2023. URL https://ideas.repec.org/p/bog/wpaper/322.html.

Z. Huang, H. Liu, and T. Wang. Modeling long memory volatility using realized measures of volatility: A realized har garch model. *Economic Modelling*, 52:812–821, 2016. ISSN 0264-9993. doi: https://doi.org/10.1016/j.econmod.2015.10.018. URL https://www.sciencedirect.com/science/article/pii/S0264999315003144.

A. J. Patton. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1):246–256, January 2011. URL https://ideas.repec.org/a/eee/econom/v160y2011i1p246-256.html.