University of Pretoria
Department of Economics Working Paper Series

A Moving Average Heterogeneous Autoregressive Model for Forecasting the Realized Volatility of the US Stock Market: Evidence from Over a Century of Data

Afees A. Salisu
Ton Duc Thang University
Rangan Gupta
University of Pretoria
Ahamuefula E. Ogbonna
University of Ibadan
Working Paper: 2019-78
November 2019

Department of Economics
University of Pretoria
0002, Pretoria
South Africa
Tel: +27 12 420 2413

A Moving Average Heterogeneous Autoregressive Model for Forecasting the Realized Volatility of the US Stock Market: Evidence from Over a Century of Data

Afees A. Salisu[1,2], Rangan Gupta[3,*] and Ahamuefula E. Ogbonna[4,5]

[1]Department for Management of Science and Technology Development, Ton Duc Thang University, Ho Chi Minh City, Vietnam.
[2]Faculty of Business Administration, Ton Duc Thang University, Ho Chi Minh City, Vietnam. Email: afees.adebare.salisu@tdtu.edu.vn.
[3]Department of Economics, University of Pretoria, Pretoria, 0002, South Africa. Email: rangan.gupta@up.ac.za.
[4] Centre for Econometric & Allied Research, University of Ibadan.
[5] Department of Statistics, University of Ibadan.
ogbonnaephraim@yahoo.com

*Corresponding author.

A Moving Average Heterogeneous Autoregressive Model for Forecasting the Realized Volatility of the US Stock Market: Evidence from Over a Century of Data

Abstract

This study forecasts the monthly realized volatility of the US stock market covering the period of February, 1885 to September, 2019 using a recently developed novel approach – a moving average heterogeneous autoregressive (MAT-HAR) model, which treats threshold as a moving average generated time varying parameter rather than as a fixed or unknown parameter. The significance of asymmetric information in realized volatility of stock market forecasting is also considered by examining the case of good and bad realized volatility. The Clark and West (2007) forecast evaluation approach is employed to evaluate the forecast performance of the proposed predictive model vis-à-vis the conventional HAR and threshold HAR (T-HAR) models. We find evidence in favour of the MAT-HAR model relative to the HAR and T-HAR models. Also observed is the significant role of asymmetry in modeling the realized volatility as good realized volatility and bad realized volatility yield dissimilar predictability results. Our results are not sensitive to the choice of sample periods and realized volatility measures.

JEL Codes: C22, C53, G12

Keywords: Realized volatility, US stock market, Forecast evaluation, HAR models

A Moving Average Heterogeneous Autoregressive Model for Forecasting the Realized Volatility of the US Stock Market: Evidence from Over a Century of Data

1.0 Introduction

Unlike the case of financial asset returns, researchers widely agreed that return volatility is predictable (Bauer and Vorkink, 2011). However, noble distinction has been made about the definition of market return volatility. This dwells on whether market volatility should be defined by fitting parametric econometric models such as generalized autoregressive conditional heteroscedasticity (GARCH) or as realized volatility. The term realized volatility was first formalized by Anderson et al. (2001) and Barndorff-Nielsen and Shephard (2002) as an alternative to the conventional market volatility measure (see Bauer and Vorkink, 2011). According to Anderson et al. (2001), who computed daily realized volatility simply by summing intraday squared returns, the process of generating realized volatility avoids possible misspecification that is possible in the conventional method, and allows returns volatility to be expressed as an observed series rather than as a latent variable as indicated by the competing parametric volatility models (see Bauer and Vorkink, 2011). The argument was supported by Barndorff-Nielsen and Shephard (2002), which defined realized volatility as the quadratic variation-like measure of activity in financial markets.

The implications of accurate forecast of realized volatility of stock market returns have been documented in the finance literature. Accurate forecast of returns volatility has been described to have important implications for financial economics and portfolio risk managers in making economic policies on asset allocation, asset pricing and risk management (see Gong and Lin, 2019). More specifically, with the prediction for financial shocks or lower returns in the developed stock market such as the US or UK stock market, investors may diversify into a less correlated stock market; mostly of developing countries such as Nigeria (Oloko, 2018). US stock market investors may also which to pursue efficient portfolio management strategies by taking short position (selling assets) or long position (buying assets) in the oil market (see Salisu and Oloko, 2015). Accurate forecasting of stock market realized volatility may also be useful for the overall economic planning and management, as stock market performance approximately represents the overall economic performance (Solnik, 1974; Salisu et al., 2019).

Due to the significant benefit of accurate forecast of realized assets volatility, a number of studies have been conducted on realized volatility in recent time. These include studies on realized volatility of US stock market such as; Bauer and Vorkink (2011), Chen et al. (2019), Gong and Liu (2019), Gupta et al. (2018), Wu and Wang (2019) and Yao et al. (2019). Other studies such as Mei et al. (2017), Ping and Li (2018), Peng et al. (2018) and Zou et al. (2019) have investigated realized volatility of Chinese stock market, while Degiannakis (2017; 2018), Sharma and Vipur (2016) investigated realized volatility of stock markets of selected multiple countries, while Salisu and Ogbonna (2018) investigated the importance of time variation in the stochastic volatility component of G7 countries. Realized volatility of other assets such as agricultural commodity future (see Tian et al., 2017; Yaug et al., 2017), oil future (see Degiannakis and Filis, 2017; Ma et al. 2017; Ma et al., 2018), electricity (see Qu et al., 2018), gold (see Gkillas et al., 2019a) and exchange rate (see Gkillas et al., 2019b).

The main objective of this study is to forecast the realized volatility of the US stock market using a novel approach. The conventional heterogeneous autoregressive (HAR) model of Corsi (2009) has been severely criticized for inefficiency due to its failure to account for some important features of financial market (see for example, Cho and Shin, 2016; Qu et al., 2018; Yau et al., 2019). Consequently, notable modifications to the model have been suggested and implemented by many researchers. For instance, Bauer and Vorkink (2011) proposed four distinct multivariate HAR models (MHAR-RV-BP, MHAR-RV-BPA, MHAR-RV-X, and MHAR-RV-BPAX) to augment realized volatility in the conventional multivariate HAR model with the principal components of bi-power covariation, asymmetry, stock returns predictors and predictors from other models, respectively. In the same vein, Cho and Shin (2016) proposed integrated HAR (IHAR) model, where the sum of the HAR coefficients is constrained to one. Chen et al. (2019) proposed HAR realized volatility model that accounts for jumps, co-jumps and their asymmetry in the predictive model. Gong and Liu (2019) accounted for leverage effect, while Yau et al. (2019) modified the HAR model by adopting hierarchical clustering technique to get a "cluster HAR".

One prominent factor that has been commonly proved to be significant in improving the forecast performance of the HAR model for realized stock market volatility is time variation. The importance of accounting for time variation in the parameters of the model has been emphasized by many studies (see for example, Bauer and Vorkink, 2011; Ma et al., 2017; Tian et al., 2017). Specifically, Bauer and Vorkink, 2011 noted that the estimated elasticities of some variables in the predictive model changes over the sample period. Ma et al. (2017) show that HAR-RV models with regime switching increase the forecasting ability significantly than those without regime switching. Tian et al. (2017) show that the proposed HAR model with time-varying sparsity improves the forecast performances substantially relative to both the simple HAR model and more sophisticated HAR-type models in almost all cases. Similarly, Qu et al. (2018) noted that accounting for time-varying volatility of realized volatility improves forecast performance of the predictive model. This is also supported by Gupta et al. (2018), Ping and Li (2018) and Wu and Wang (2019).

This study contributes to the extant literature on forecasting realized volatility with heterogeneous autoregressive (HAR) model by accounting for threshold effect in a unique way with the adoption of a moving average threshold heterogeneous (MAT-HAR) model. Our proposed model, as developed by Motegi et al., (2019), is a novel combination of HAR and threshold autoregressive (TAR) models. As in the HAR model, the MAT-HAR model has multiple groups of lags of a target series, where the groups are constructed from a viewpoint of sampling frequencies. Tapping from the TAR properties also, this method allows for the presence of a threshold term in each group. The threshold is an observed moving average of lagged target series, which guarantee time-varying thresholds and simple estimation with least squares estimator. As evident from the Monte Carlo simulation conducted by Motegi et al. (2019), MAT-HAR was found to have higher forecast accuracy than the HAR baseline model.

Earlier studies on the realized volatility of US stock market such as Bauer and Vorkink (2011), Chen et al. (2019), Gong and Liu (2019), Gupta et al. (2018), Wu and Wang (2019) and Yao et al. (2019) have not considered moving average threshold effect as a modification to the conventional HAR model. While recent studies such as Gupta et al. (2018) and Wu and Wang (2019) appear to account for time variation in the predictive model, they assume unknown threshold; which complicates both estimation and inferences (Motegi et al., 2019). Meanwhile, following Chen et al. (2019), Qu et al., (2017) and Wu and Wang (2019) that find leverage effect to be significant in forecasting realized volatility, this study accounts for realized return volatility asymmetry by distinguishing between the good realized volatility and bad realized volatility. Our proposed model is expected to outperform the conventional HAR model and HAR model with non-moving average threshold. Whereas, MAT-HAR model with asymmetry may be expected to outperform symmetric MAT-HAR model if asymmetry is significant in the model (see Salisu and Isah, 2017).

In addition to the usage of the MAT-HAR model for the first time in forecasting the realized volatility of the US stock market, another unique feature of our paper is the usage of the longest available historical data for the Dow Jones Industrial Average (DJIA), covering the period of February, 1885 to September, 2019. In the process, our analysis is likely to be robust to the problem associated with the sample selection bias, as we track the entire history of the DJIA since its inception to the current period. Following this introductory section, the rest of the paper is structured as follows. The model is presented in Section 2. Section 3 deals with presentation and discussion of results. Section 4 concludes the paper.

2.0 The Model

Basically, this study proposes the usage of the moving average threshold heterogeneous autoregressive (MAT-HAR) model of Motegi et al., (2019), which is a hybrid of both heterogeneous autoregressive (HAR) model by Corsi (2009) and the threshold autoregressive model by Tong (1978). The main problem of the HAR model is that threshold effects are not considered (see Corsi et al., 2012). This assumption may however be unrealistic in the present world where financial market and macroeconomic indicators have been found to have several regimes that switch stochastically over time (Salisu et al., 2019).

On the other hand, the TAR-type model considers threshold effect but assumes that threshold is unknown or constant over time. The problem associated with this is that the number of lags included in the model needs to be small in order to keep estimation and inference simple. Furthermore, the assumption of constant threshold can be unrealistic, given that threshold may be vary over time based on the state of the economy. In essence, MAT-HAR model has multiple groups of lags of a target series, where groups are constructed from a viewpoint of sampling frequencies as in the HAR model, and allows for the presence of a threshold term in each group as in the TAR model. However, the threshold is treated as an observed moving average of lagged target series, which guarantees time-varying thresholds and simple estimation through the method of least squares.

The specification of MAT-HAR model is started by first specifying the HAR model. Assuming $\{y\}_{t=1}^{n}$ is the realized (overall, good or bad) volatility series, a HAR model can be specified as:

$$y_t = \beta^{(0)} + \sum_{k=1}^{K} \beta^{(k)} y_{t-1}^{(k)} + u_{t,} \tag{1}$$

where

$$y_t^{(k)} = \frac{1}{m_k} \sum_{i=1}^{m_k} y_{t+1-i,} \tag{2}$$

$\beta = (\beta^{(0)}, \beta^{(1)}, ..., \beta^{(K)})'$ are parameters to be estimated, u_t is an error term, and $(m_1, m_2, ..., m_k)$ are pre-specified values that represent sampling frequencies. In the context of our study where we are dealing with monthly data frequency, $(m_1, m_2, m_3) = (1, 3, 12)$; representing one month, one quarter and one year, respectively.

The second part of MAT-HAR model is the TAR model. This can be specified as follows.

$$y_t = \beta_0 + \beta_1 y_{t-1} + \eta_1 1(y_{t-1} > \mu) y_{t-1} + u_t, \tag{3}$$

where μ is an unknown threshold to be estimated and $\mathbf{1}(y_{t-1} > \mu)$ is an indicator function that equals 1 if $y_{t-1} \geq \mu$ and 0 otherwise. A naïve way to add threshold terms to the HAR model would be to specify as:

$$y_t = \beta^{(0)} + \sum_{k=1}^{K} \left\{ \beta^{(k)} y_{t-1}^{(k)} + \eta^{(k)} 1(y_{t-1}^{(k)} \geq \mu_k) y_{t-1}^{(k)} \right\} + u_t, \tag{4}$$

In equation (4), thresholds terms are added to each of the K sampling frequencies. The two problems with this model are that; (i) the presence of K unknown thresholds, $(\mu_1, \mu_2, ..., \mu_K)$, complicates estimation and inference substantially. The second problem is that thresholds are fixed over time, which could be an unrealistic assumption.

We, as in Motegi et al., (2019), circumvent the above mentioned problems by specifying the thresholds to be the moving averages of the lagged target variable instead of specifying them to be mere constants. The proposed MAT-HAR model is thus specified as:

$$y_t = \beta^{(0)} + \sum_{k=1}^{K} \left\{ \beta^{(k)} y_{t-1}^{(k)} + \eta^{(k)} 1(y_{t-1}^{(k)} \geq \hat{\mu}_{t-1}^{(k)}) y_{t-1}^{(k)} \right\} + u_t, \tag{5}$$

where

$$\hat{\mu}_t^k = \frac{1}{t+1-\max\left\{t-\ell_n, 1\right\}} \sum_{\tau=\max\{t-\ell_n,1\}}^{t} y_\tau^{(k)} \tag{6}$$

$\hat{\mu}_t^k$ is the moving average of $y_t^{(k)}$ using ℓ_n lags (or all lags available if ℓ_n>t). The MAT-HAR model (5) resolves the two issues on model (4); first, by ensuring that $\hat{\mu}_t^k$ and hence $\mathbf{1}(y_{t-1}^{(k)} \geq \hat{\mu}_{t-1}^{(k)}$ can be computed from data. Hence, $\theta = \left(\beta^{(0)}, \beta^{(1)}, ..., \beta^{(K)}, \eta^{(1)}, ..., \eta^{(K)}\right)'$ can easily be estimated using OLS. Second, by ensuring that the thresholds $\hat{\mu}_t^k$ are time dependent and adopting the moving average structure that is intuitively easy to understand.

2.1 Modeling and evaluation of forecast performance

For forecast evaluation, the Clark and West (2007) test (hereafter, C-W test) is employed. Unlike the Campbell and Thompson (2008) test, the C-W test is able to determine whether the difference between the forecast errors of any two competing models is statistically significant, and also suitable for nested models. The underlying procedure for the C-W test involves calculating equation (7):

$$\hat{f}_{t+k} = \left(r_{t+k} - \hat{r}_{1t,t+k}\right)^2 - \left[\left(r_{t+k} - \hat{r}_{2t,t+k}\right)^2 - \left(\hat{r}_{1t,t+k} - \hat{r}_{2t,t+k}\right)^2\right] \tag{7}$$

where k is the forecast period, $\left(r_{t+k} - \hat{r}_{1t,t+k}\right)^2$ is the squared error for the restricted model (i.e., model 1), $\left(r_{t+k} - \hat{r}_{2t,t+k}\right)^2$ is the squared error for the unrestricted model (model 2), and $\left(\hat{r}_{1t,t+k} - \hat{r}_{2t,t+k}\right)^2$ is the adjusted squared error introduced by the C-W test to correct for any noise associated with the larger model's forecast. Thus, the sample average of $\hat{f}_{t+k}$ can be expressed as $MSE_1 - \left(MSE_2 - \text{adj.}\right)$, and each term is computed as

$$MSE_1 = P^{-1}\sum\left(r_{t+k} - \hat{r}_{1t,t+k}\right)^2 ,$$

$$MSE_2 = P^{-1}\sum\left(r_{t+k} - \hat{r}_{2t,t+k}\right)^2 \text{, and}$$

$$\text{adj.}=P^{-1}\sum\left(\hat{r}_{1t,t+k} - \hat{r}_{2t,t+k}\right)^2 ,$$

where P is the number of predictions used in computing these averages. To test for equality of forecast performance between models 1 and 2, the $\hat{f}_{t+k}$ is regressed on a constant, and the resulting t-statistic for a zero coefficient is used to draw the inference. Since the null hypothesis tests for equality of MSEs, the alternative hypothesis implies otherwise. The null hypothesis is rejected if the test statistic is greater than +1.282 (for a one-sided 0.10 test) or +1.645 (for a one-sided 0.05 test).

In this study, we hypothesize that our proposed MAT-HAR model (equation 5) will outperform the conventional HAR model (equation 1) and the HAR model with naïve threshold modeling (T-HAR) (equation 4). Thus, we compare HAR model with T-HAR and T-HAR model with MAT-HAR model. In the HAR vs. T-HAR model, HAR is the model 1 (restricted model) while T-HAR is the model 2 (unrestricted model). A relatively lower root mean square error (RMSE) for T-HAR than HAR is required to conclude than T-HAR outperforms HAR model. In the same vein, a relatively lower root mean square error (RMSE) for MAT-HAR relative to T-HAR is required to conclude that MAT-HAR outperforms T-HAR model.

Also, there exists a growing literature on the significance of asymmetric realized volatility and leverage effect in modeling and forecasting financial market performance. For instance, Alam et al. (2019) found that bad volatility dominates good volatility in terms of the importance of transmission of volatility shocks. Bensaida (2018) find evidence suggesting that the good and bad volatilities are transmitted with different time-varying intensities. He et al. (2019) find that bad volatility reacts more strongly to shocks in EPU following the debt crisis and trade negotiations. In addition, the empirical result by Gong and Lin (2017) show that bad and good uncertainties have long memory property, and the predictability of long-term good and bad uncertainties is stronger than that of short- and mid-term good and bad uncertainties (see also, Apergis et al., 2017; Barunik et al., 2016; Wu and Hou, 2019). More importantly, as Salisu and Isah (2017) noted that failure to account for asymmetry when it is significant may lead to biased result, we consider decomposition of realized volatility into two; bad realized volatility and good realized volatility. In terms of forecasting evaluation, the forecast performance of bad realized volatility is expected to be dissimilar with that of good realized volatility if asymmetry really matter. Otherwise, asymmetry does not matter when forecasting realized volatility for US stock market.

3. Data Preliminaries

The data used in this study is for the realized volatility of the DJOA spanning over a century, from February 1885 and September 2019. The underlying daily data was extracted from the database of MeasuringWorth.[1] Note that, we first compute daily log-returns (i.e., the first-difference of the natural logarithm of the daily DJIA index), and then we derive the realized volatility by summing the squared returns over the number of daily observations available for a specific month. Also, the realized volatility was disaggregated into good and bad realized volatility, with the aim of examining the potential of asymmetry to influence the prediction of the realized volatility. For good and bad volatility, we follow the same procedure described above for the derivation of the overall realized volatility, but now we concentrate on daily positive and negative returns respectively. Consequently, there are three variables to be considered – realized volatility, bad realized volatility (to reflect the uncertainty that is associated with movements in stock returns volatility that are associated with bad news) and good realized volatility (to reflect the movements in stock returns volatility that are associated with bad news), with each having a total of about 1,617 observation points. A brief summary statistics of these three variables is presented in Table 1, showing the measures of location (mean and median), spread (minimum, maximum and standard deviation) and shape (skewness, kurtosis and Jarque-Bera). All three volatility measures are found to be positively skewed (confirming why the mean and median are not equal, as with symmetric series), and as expected for returns series, they are all leptokurtic, which jointly indicate the non-normality of the realized volatility measures. Figures 1 to 3 in the Appendix show how the computed thresholds for the three different sampling frequencies trace out the corresponding sampling frequency realized volatility, which suggest their ability to track the realized volatility.

[1] The data is downloadable from: https://www.measuringworth.com/datasets/DJA/index.php.

Table 1: Summary Statistics

	Realized Volatility	Good Realized Volatility	Bad Realized Volatility
Mean	0.0028	0.0013	0.0014
Median	0.0014	0.0007	0.0006
Minimum	0.0001	0.0000	0.0000
Maximum	0.0972	0.0451	0.0802
Standard Deviation	0.0061	0.0029	0.0036
Skewness	8.8121	8.2797	11.0245
Kurtosis	107.3069	92.1099	182.8744
Jarque-Bera Statistic	753963.20***	553470.50***	2212659.00***
Observations	1617	1617	1617

Note: *** indicates statistical significance at 1% level.

4. Empirical Results

In this study, we aim to investigate the forecast performance of the US stock market volatility using a novel approach; the moving average threshold heteroscedastic autoregressive (MAT-HAR) model.[2] As noted earlier, the model extends the conventional heteroscedastic autoregressive (HAR) model by combining a HAR model with a threshold autoregressive (TAR) model, in order to account for plausible dynamics of threshold effects. Consequently, we begin from the baseline model, which is the HAR model that does not incorporate any form of threshold effect, and then, consider two other variants of the baseline model that incorporate threshold parameters. Therefore, the two variants differ only by the assumptions relating upon which the threshold parameters are incorporated. First, a single threshold value is used to capture the threshold effect that may be inherent in the series; thus, giving rise to an extended HAR variant – the threshold HAR model (T-HAR). This variant usually involves very tasking numeric procedure, in a bid to determine an optimal threshold. For the purpose of our empirical study, we consider that stance where the average of the series is adopted as the threshold.

Second, attempting to circumvent the difficulties associated with the numerical search for an optimal threshold, the moving average framework is introduced to capture the threshold effects. This is a direct deviation from the use of one single value as a threshold to accommodating the fact that the thresholds may not be static, but dynamic. Thus, for every new data point, a new threshold

[2] All estimations were conducted in MATLAB using the codes available on the website of Professor Kaiji Motegi at: http://www2.kobe-u.ac.jp/~motegi/research.html. We are indeed thankful to Professor Motegi for making his codes publicly available.

is computed by averaging over all the observation points, iteratively. This variant is the novel HAR approach – the MAT-HAR model, which attempts to account for an inherent salient feature (see Salisu and Ogbonna, 2019) of time variation, with respect to thresholds. Consequently, this study will examine the forecast performance of three different models – HAR model (baseline model), T-HAR and MAT-HAR. Also, given that these models are nested the forecast performances are examined with Clark and West (2007) statistic, in addition to the conventional forecast error measure – RMSE, while attempting to answer two main questions. First, "Does threshold matter in forecasting the realized volatility of stock market with HAR-based model?" In this, the T-HAR and MAT-models are compared with the baseline model – HAR model, to ascertain whether, or not, the HAR model variants that incorporate threshold parameters (either as a single value or as a moving average) would outperform the baseline HAR model that ignores same. Second, "Does the MAT-HAR outperform the HAR and T-HAR models?" Third, "Does asymmetry matter in the MAT-HAR?" In consonance with the HAR framework, which specifies three different sampling frequencies (monthly $(m=1)$, quarterly $(m=3)$ and yearly $(m=12)$) for the realized volatility series, three threshold variables corresponding to each sampling frequency are incorporated singly and/or simultaneously.

Following from the above, we estimate a total of eight different models *(Model1 – Model8)* in this study. The first model is our baseline model – the HAR model, in which threshold effect is assumed not to exist at any of the sampling frequencies. Model2, Model3 and Model4 specify the threshold variants that assume the existence of threshold effect (using either constant or time varying threshold variables) at frequencies 1, 2 and 3, respectively. Model5, Model6 and Model7 assume that the threshold effect may be existent in two different sampling frequencies and consequently, correspond to threshold effects existing at sampling frequencies 1 and 2; frequencies 1 and 3; and frequencies 2 and 3; respectively. The last model – Model8 assumes that threshold effect exists at all three frequencies. Across the last seven models, the constancy or time varying nature of the threshold variable is examined for all three realized volatility measures. We therefore proceed to examine the forecast error of all eight models using the RMSE and Clark and West statistic for the in-sample period (ranging over periods between January, 1901 and September, 2017) and three out-of-sample periods – $6, 12$ and 24 months ahead forecasts.

Table 2: In-sample RMSE Result

Model	T-HAR			MAT-HAR		
	Realized Volatility	***Bad Realized Volatility***	***Good Realized Volatility***	***Realized Volatility***	***Bad Realized Volatility***	***Good Realized Volatility***
Model1	0.005054961	0.003386985	0.002292152	0.005054961	0.003386985	0.002292152
Model2	0.005054961	0.003386770	0.002291942	0.005052049	0.003386950	0.002292064
Model3	0.005053126	0.003386835	0.002292104	0.005054878	0.003386468	0.002292129
Model4	0.005054282	0.003386774	0.002292146	0.005044100	0.003386943	0.002289595
Model5	0.005052869	0.003386434	0.002291939	0.005051507	0.003386252	0.002292064
Model6	0.005054272	0.003386547	0.002291919	0.005035952	0.003386927	0.002289594
Model7	0.005051674	0.003386672	0.002292083	0.005039984	0.003386034	0.002289439
Model8	0.005051465	0.003386286	0.002291909	0.005035324	0.003385919	0.002289403

On the in-sample RMSE results for realized volatility (see Table 2) when the T-HAR model was used, we find relatively small differences across all eight models, with *Model1* and *Model2* not markedly different even at the tenth decimal place. However, all other models seem to have smaller RMSEs that the first two models, in the in-sample period. The case is different when MAT-HAR was used to predict the in-sample realized volatility, as all the models incorporating time varying thresholds had lower RMSEs than Model1 that does not. This result is consistent with the findings by Gupta et al. (2018), Ma et al. (2017), Ping and Li (2018), Qu et al., (2017), Tian et al. (2017) and Vorkink (2011), which find evidence of significant time varying effects.

With regards to the Bad and Good realized volatilities, all the models incorporating thresholds (either constant or time varying) were found to have lower forecast error than the baseline model that neglects the threshold variable. Although, relatively small differences, models incorporating threshold variables appears to generally forecast the three different realized volatility measures with smaller forecast errors than the baseline model. Further probing the differences in the RMSEs of the contending models in the out-of-sample (see Table 3) periods, we find the similar stances as with the in-sample period except for the 6-months ahead forecast horizon for the T-HAR model, under *Model1* and *Model2*, with the former having lower RMSE than the latter. However, again the models with threshold variables seemed to consistently have lower RMSEs than those the failed to incorporate threshold variables at any frequency. However, the suggested outperformance of the baseline model by the one or all of the seven models incorporating threshold variables at the one or all of the three specified frequencies, is further subjected to the pair-wise comparison using the Clark and West (2007) statistic (see Tables 4 and 5 for the in-sample and out-of-sample periods, respectively), which provides a more formal basis to draw conclusions.

Table 3: Out-of-sample RMSE Result

Model	T-HAR			MAT-HAR		
	$h=6$	$h=12$	$h=24$	$h=6$	$h=12$	$h=24$
			Realized Volatility			
Model1	0.005044578	0.005034213	0.005012818	0.005044578	0.005034213	0.005012818
Model2	0.005044579	0.005034212	0.005012828	0.005041643	0.005031210	0.005009787
Model3	0.005042822	0.005032441	0.005011228	0.005044492	0.005034122	0.005012723
Model4	0.005043948	0.005033651	0.005012226	0.005033826	0.005023657	0.005002419
Model5	0.005042578	0.005032230	0.005010982	0.005041104	0.005030669	0.005009257
Model6	0.005043941	0.005033638	0.005012231	0.005025647	0.005015389	0.004994136
Model7	0.005041456	0.005031180	0.005009940	0.005029694	0.005019553	0.004998315
Model8	0.005041263	0.005031015	0.005009745	0.005025016	0.005014782	0.004993532
			Bad Realized Volatility			
Model1	0.003380135	0.003373136	0.003358882	0.003380135	0.003373136	0.003358882
Model2	0.003379940	0.003372933	0.003358640	0.003380098	0.003373093	0.003358833
Model3	0.003380013	0.003373003	0.003358798	0.003379626	0.003372637	0.003358409
Model4	0.003379904	0.003372872	0.003358645	0.003380092	0.003373088	0.003358829
Model5	0.003379661	0.003372636	0.003358391	0.003379405	0.003372401	0.003358163
Model6	0.003379698	0.003372659	0.003358390	0.003380075	0.003373068	0.003358806
Model7	0.003379829	0.003372792	0.003358603	0.003379188	0.003372183	0.003357944
Model8	0.003379491	0.003372443	0.003358213	0.003379070	0.003372056	0.003357811
			Good Realized Volatility			
Model1	0.002287352	0.002282571	0.002272892	0.002287352	0.002282571	0.002272892
Model2	0.002287117	0.002282328	0.002272662	0.002287272	0.002282494	0.002272817
Model3	0.002287306	0.002282525	0.002272857	0.002287332	0.002282551	0.002272872
Model4	0.002287350	0.002282571	0.002272891	0.002284836	0.002280121	0.002270466
Model5	0.002287117	0.002282328	0.002272665	0.002287271	0.002282494	0.002272818
Model6	0.002287097	0.002282315	0.002272646	0.002284835	0.002280119	0.002270464
Model7	0.002287290	0.002282516	0.002272846	0.002284675	0.002279966	0.002270311
Model8	0.002287092	0.002282311	0.002272646	0.002284643	0.002279937	0.002270284

4.1. Does threshold matter in the HAR-based realized volatility of stock market?

Here, we comparatively examine the forecast performance of seven pairs of models, comprising of the baseline model *(Model1)* and each of the other contending models *(Model2 – Model8)*, using a more formal approach – the Clark and West statistic. The aim is to ascertain the relevance, or otherwise, of the incorporating one or more threshold variables in the HAR model framework for predicting different realized volatility measures. Consequently, we progress the comparison in a sequential order: First, comparing T-HAR model (incorporating threshold variables that are constant over time *(Model2 – Model8)*) with the baseline model (*Model1*) and ascertain the model that has the better predictive performance. Consequently, an outperformance of T-HAR model over the baseline HAR model would suggest that threshold variable(s) that is constant over time does matter in the prediction of the realized volatility. However, if the reverse is the case, then incorporating a threshold variable that is constant over time may not yield any better results than the conventional HAR model. Second, the better model from the first comparison would then be compared with the MAT-HAR model. Similarly, an outperformance of the MAT-HAR model over the better model from the first comparison would suggest the importance of not just incorporating a threshold variable, but ensuring that it is one that is time varying, as it is more likely to accommodate the inherent dynamics. These comparisons are carried out both for the in-sample period and all three out-of-sample periods, across the three measures of realized volatilities and all the contending models.

4.2 Does the T-HAR outperform the HAR?

In a bid to examine the forecast outperformance stance between the conventional HAR model (*Model1*) and the T-HAR model variants (*Model2 – Model8*), we adopt the Clark and West test statistics, which formally compares a pair of nested models – one restricted (here, the HAR model) and one unrestricted model (here, T-HAR model), under the null hypothesis that the restricted model is preferred over the unrestricted model. Therefore, positive and statistically significant Clark and West test statistic would indicate preference for the corresponding T-HAR variant over the conventional HAR model, while non-significance would imply an outperformance of the latter over the former. The in-sample and out-of-sample results are presented in Tables 4 and 5, respectively; under the assumption of time invariant threshold variable. The various T-HAR model constructs (*Model2 – Model8*) do not seem to significantly outperform the conventional HAR model (*Model1*), in predicting the realized volatility in the in-sample period. This feat is further replicated across the three out-of-sample forecast horizons used in this study. Also worthy of note, the threshold effects that the time invariant threshold HAR model constructs presume to account for, regardless of sampling frequencies (ranging from 1 - 3) for which they are incorporated, seem not to significantly improve upon the conventional HAR model's performance. Put differently, a HAR model that incorporates a time invariant threshold variable is less likely to outperform the conventional HAR model that does not account for any threshold parameters. This is possibly an immediate fall out of the rather stringent and restrictive stance of assuming that the threshold variables are time-invariant, which fails to take into cognizance the dynamic nature of volatility series. The single threshold value does not seem to provide sufficient information to distinguish the performance of the T-HAR model variants from the conventional HAR model. From the above stance, accounting for threshold effects may not matter, if the threshold value is constant over time.

4.3 Does the MAT-HAR outperform the conventional HAR?

Following from the results of no marked difference between the conventional HAR model and the T-HAR model variants that incorporate threshold variables at different sampling frequencies, using single values, we proceed to compare the conventional HAR model (*Model1*) with the HAR model variants (*Model2 – model8*) that allow for more dynamic threshold variables. Similar to the earlier comparison, the Clark and West (2007) test statistics is used to formally compare different pairs of nested models – a restricted model (here, the HAR model) and an unrestricted model (here, MAT-HAR model). The null hypothesis that the restricted model is preferred over the unrestricted model, is rejected whenever a positive and statistically significant Clark and West test value is obtained. The intuition here is to ascertain whether the non-performance of the T-HAR model variants (*Model2 – Model8*) over the conventional HAR model could be salvaged by allowing for a time varying threshold variable, in a bid to validate the relevance, or otherwise, of accounting for threshold effects in the heteroscedastic autoregressive model framework. The results of the in-sample and out-of-sample pairwise comparisons of the conventional HAR model (*Model1*) and the MAT-HAR variants (*Model2 – Model8*) are presented in Tables 4 and 5, respectively, under the assumption that the threshold variable is time varying. Consistent with the Clark and West (2007) test statistic, positively significant values would indicate preference for the corresponding

MAT-HAR variant over the conventional HAR model (*Model1*). In contrast to the stance when time-invariant threshold variables were incorporated in the T-HAR framework, a markedly different stance is here obtained.

On the in-sample period, we find outperformance of four MAT-HAR variants (*Model4, Model6, Model7* and *Model8*) over the conventional HAR model (*Model1*), while the predictive performance of the remaining three MAT-HAR model variants (*Model2*, *Model3* and *Model5*) were not statistically distinguishable from the conventional HAR model. The results suggests that HAR model variants that incorporate time-varying threshold variables; singly (at yearly sampling frequency (*Model4*)) or in pairs (at the monthly and yearly (*Model6*), and at the quarterly and yearly (*Model7*) sampling frequencies) or all three sampling frequencies (*Model8*); appear to outperform the conventional HAR model in the in-sample. *Model2*, *Model3* and *Model5* do not perform better than the conventional HAR model, in the prediction of realized volatility. Similar feats of outperformances are replicated across the three out-of-sample forecast horizons considered. The stances of the consistent outperformance the four MAT-HAR variants (*Model4, Model6, Model7* and *Model8*) both in the in-sample and out-of-sample forecast period suggest that the models are robust to the choice of sample periods.

However, given that some MAT-HAR variants do not outperform the conventional HAR model, both in the in-sample and out-of-sample periods, the MAT-HAR model predictive performance may be dependent on the sampling frequencies at which the threshold effects are incorporated. Generally, accounting for inherent salient statistical features, as opined by Salisu and Ogbonna (2019), by allowing for time-varying thresholds (following Motegi et al. (2019) approach) does matter when predicting the realized volatilities, as it is found to improve models' predictive performance (see Qu et al., 2017; Gupta et al., 2018; Ping and Li, 2018; Motegi et al., 2019; and Wu and Wang, 2019; among others). Therefore, incorporating time-varying thresholds at the significant sampling frequencies have higher tendencies to improve the forecast performance of the HAR model than ignoring accounting for these thresholds and/or assuming thresholds that are constant over time. Consequent upon these feats, an optimal MAT-HAR model would be determined from among the contending MAT-HAR model constructs.

Table 4: In-sample Clark and West Result

Models	T-HAR			MAT-HAR		
	Realized Volatility	***Bad Realized Volatility***	***Good Realized Volatility***	***Realized Volatility***	***Bad Realized Volatility***	***Good Realized Volatility***
Model2	5.46E-14 [7.10E-11]	2.91E-09 [1.26E-08]	1.92E-09 [4.19E-09]	5.89E-08 [7.43E-08]	4.73E-10 [4.08E-09]	8.06E-10 [3.68E-09]
Model3	3.71E-08 [4.13E-08]	2.04E-09 [5.96E-09]	4.35E-10 [2.20E-09]	1.68E-09 [1.11E-08]	7.01E-09 [1.23E-08]	2.10E-10 [2.16E-09]
Model4	1.37E-08 [2.42E-08]	2.86E-09 [5.91E-09]	4.81E-11 [5.89E-10]	2.19E-07*** [7.61E-08]	5.75E-10 [4.00E-09]	2.34E-08* [1.26E-08]
Model5	4.23E-08 [3.04E-08]	7.47E-09 [1.24E-08]	1.95E-09 [4.48E-09]	6.98E-08 [7.79E-08]	9.93E-09 [1.38E-08]	8.06E-10 [3.64E-09]
Model6	1.39E-08 [2.28E-08]	5.94E-09 [1.16E-08]	2.13E-09 [4.46E-09]	3.84E-07*** [1.44E-07]	7.86E-10 [5.33E-09]	2.34E-08* [1.25E-08]
Model7	6.64E-08 [5.09E-08]	4.24E-09 [8.30E-09]	6.27E-10 [2.90E-09]	3.02E-07*** [1.05E-07]	1.29E-08 [1.51E-08]	2.49E-08* [1.36E-08]
Model8	7.07E-08 [4.65E-08]	9.48E-09 [1.14E-08]	2.22E-09 [5.09E-09]	3.96E-07*** [1.48E-07]	1.44E-08 [1.69E-08]	2.52E-08* [1.40E-08]

Note: The listed contending models are compared with *Model1* (the conventional HAR model), which is the baseline model. The figures in square brackets are the associated standard errors of the estimated statistics. ***, ** and * denote significance at 1%, 5% and 10% levels, respectively. Statistical significance of the Clark and West (2007) test indicates preference for the contending model over the baseline model.

Table 5: Out-of-sample Clark and West Result

Model	T-HAR			MAT-HAR		
	$h=6$	$h=12$	$h=24$	$h=6$	$h=12$	$h=24$
Realized Volatility						
Model2	-8.37E-12 [7.22E-11]	9.33E-12 [7.66E-11]	-9.87E-11 [8.30E-11]	5.89E-08 [7.40E-08]	5.95E-08 [7.37E-08]	5.94E-08 [7.30E-08]
Model3	3.64E-08 [4.11E-08]	3.65E-08 [4.10E-08]	3.46E-08 [4.06E-08]	1.71E-09 [1.10E-08]	1.76E-09 [1.10E-08]	1.78E-09 [1.09E-08]
Model4	1.33E-08 [2.41E-08]	1.26E-08 [2.40E-08]	1.28E-08 [2.38E-08]	2.18E-07*** [7.58E-08]	2.15E-07*** [7.55E-08]	2.13E-07*** [7.49E-08]
Model5	4.16E-08 [3.03E-08]	4.14E-08 [3.02E-08]	3.97E-08 [2.99E-08]	6.98E-08 [7.76E-08]	7.04E-08 [7.72E-08]	7.01E-08 [7.66E-08]
Model6	1.34E-08 [2.27E-08]	1.28E-08 [2.26E-08]	1.29E-08 [2.25E-08]	3.82E-07*** [1.44E-07]	3.80E-07*** [1.43E-07]	3.76E-07*** [1.42E-07]
Model7	6.49E-08 [5.06E-08]	6.39E-08 [5.04E-08]	6.20E-08 [5.00E-08]	3.01E-07*** [1.04E-07]	2.98E-07*** [1.04E-07]	2.95E-07*** [1.03E-07]
Model8	6.91E-08 [4.64E-08]	6.79E-08 [4.62E-08]	6.62E-08 [4.58E-08]	3.94E-07*** [1.48E-07]	3.92E-07*** [1.47E-07]	3.88E-07*** [1.46E-07]
Bad Realized Volatility						
Model2	2.77E-09 [1.26E-08]	2.82E-09 [1.25E-08]	3.08E-09 [1.24E-08]	4.87E-10 [4.06E-09]	5.26E-10 [4.04E-09]	5.62E-10 [4.01E-09]
Model3	1.85E-09 [5.93E-09]	1.92E-09 [5.91E-09]	1.59E-09 [5.86E-09]	6.94E-09 [1.22E-08]	6.85E-09 [1.22E-08]	6.64E-09 [1.20E-08]
Model4	3.01E-09 [5.89E-09]	3.24E-09 [5.87E-09]	3.05E-09 [5.82E-09]	5.81E-10 [3.98E-09]	6.09E-10 [3.97E-09]	6.41E-10 [3.93E-09]
Model5	6.97E-09 [1.24E-08]	7.15E-09 [1.23E-08]	7.05E-09 [1.22E-08]	9.88E-09 [1.37E-08]	9.89E-09 [1.37E-08]	9.72E-09 [1.36E-08]
Model6	5.94E-09 [1.16E-08]	6.23E-09 [1.15E-08]	6.30E-09 [1.14E-08]	8.00E-10 [5.31E-09]	8.50E-10 [5.29E-09]	9.01E-10 [5.24E-09]
Model7	4.20E-09 [8.27E-09]	4.46E-09 [8.23E-09]	4.03E-09 [8.16E-09]	1.28E-08 [1.51E-08]	1.28E-08 [1.50E-08]	1.27E-08 [1.49E-08]
Model8	9.13E-09 [1.13E-08]	9.49E-09 [1.13E-08]	9.28E-09 [1.12E-08]	1.44E-08 [1.68E-08]	1.45E-08 [1.67E-08]	1.43E-08 [1.66E-08]
Good Realized Volatility						
Model2	2.05E-09 [4.17E-09]	2.08E-09 [4.15E-09]	2.01E-09 [4.12E-09]	7.71E-10 [3.67E-09]	7.52E-10 [3.65E-09]	7.35E-10 [3.62E-09]
Model3	4.30E-10 [2.19E-09]	4.30E-10 [2.18E-09]	3.77E-10 [2.16E-09]	1.99E-10 [2.15E-09]	1.99E-10 [2.14E-09]	1.94E-10 [2.12E-09]
Model4	3.66E-11 [5.86E-10]	2.40E-11 [5.84E-10]	2.64E-11 [5.79E-10]	2.32E-08* [1.25E-08]	2.29E-08* [1.25E-08]	2.26E-08* [1.23E-08]
Model5	2.06E-09 [4.46E-09]	2.09E-09 [4.44E-09]	2.01E-09 [4.40E-09]	7.72E-10 [3.63E-09]	7.52E-10 [3.61E-09]	7.33E-10 [3.58E-09]
Model6	2.24E-09 [4.44E-09]	2.24E-09 [4.42E-09]	2.18E-09 [4.38E-09]	2.32E-08* [1.24E-08]	2.29E-08* [1.24E-08]	2.26E-08* [1.23E-08]
Model7	5.98E-10 [2.89E-09]	5.67E-10 [2.88E-09]	5.19E-10 [2.85E-09]	2.46E-08* [1.35E-08]	2.43E-08* [1.35E-08]	2.40E-08* [1.33E-08]
Model8	2.31E-09 [5.07E-09]	2.30E-09 [5.04E-09]	2.23E-09 [5.00E-09]	2.49E-08* [1.39E-08]	2.46E-08* [1.39E-08]	2.43E-08* [1.38E-08]

Note: The listed contending models are compared with *Model1* (the conventional HAR model), which is the baseline model. The figures in square brackets are the associated standard errors of the estimated statistics. ***, ** and * denote significance at 1%, 5% and 10% levels, respectively. Statistical significance of the Clark and West (2007) test indicates preference for the contending model over the baseline model.

4.4 Does asymmetry matter in the MAT-HAR?

Having established the relevance of allowing for time-varying, rather than time-invariant, thresholds, we further examine the variants of the MAT-HAR model on the basis of their performances when plausible asymmetry in the realized volatility is considered. This in tandem with extant researches that suggest accounting for plausible asymmetry when forecasting economic and financial series, especially, those relating to stock returns (see Bauer and Vorkink, 2011; Salisu and Isah, 2017; Salisu and Ogbonna, 2018; Chen et al., 2019; Salisu and Ogbonna, 2019; among others). Therefore, in addition to ascertaining the forecast performance of the MAT-HAR model variants in predicting realized volatility, we also consider the cases where asymmetry in realized volatility is captured either as bad realized volatility (uncertainty that is associated with negative stock prices/returns) or good realized volatility (uncertainty that is associated with positive stock prices/returns). Similar to the procedure that has previously been followed, the conventional HAR model and MAT-HAR model variants are also used to predict separately, bad and good realized volatilities. In a pair-wise comparison framework of the Clark and West (2007), the forecast performance of the models using bad and good realized volatilities are examined (see results in Tables 4 and 5). Also, in another strand of comparison, the MAT-HAR model with bad realized volatility is compared with those of the good realized volatility (see Table 6).

From Tables 4 and 5, which corresponds to the cases of the in-sample and out-of-sample periods, all the MAT-HAR model variants do not outperform the conventional HAR model in the prediction of bad realized volatility, regardless of the thresholds associated with the sampling frequencies. However, when the good realized volatility was considered, we found outperformance for four MAT-HAR model variants (*Model4, Model6, Model7 and Model8*). These models' outperformance over the conventional HAR model is consistent with the stance of their outperformance when realized volatility (not disaggregated) was the variable to be predicted. This indicates the MAT-HAR model ability to predict realized volatility and good realized volatility, but not bad realized volatility. In order to confirm this further, the second strand of comparison was conducted. This was to ascertain if the MAT-HAR model variants' predict bad and good realized volatilities in the same manner. The Clark and West (2007) result is presented in Table 6, for the in-sample and out-of-sample periods. Except for *Model2*, all other MAT-HAR model variants (*Model3* – *Model8*) are observed to predict bad and good realized volatilities quite differently. The feat of outperformance of the MAT-HAR variants in predicting good realized volatility over similar MAT-HAR variants used in predicting bad realized volatility is consistent across both in-sample and out-of-sample periods. For MAT-HAR model incorporating threshold effect only at the monthly frequency, bad and good realized volatilities are predicted in much the same way. Evidently, asymmetry seems to matter in the prediction of realized volatility, as well as the sampling frequency for which a corresponding threshold effect is incorporated.

Table 6: Clark and West Result on Asymmetry using MAT-HAR model variants

	In-Sample	$h=6$	$h=12$	$h=24$
Model2	7.68E-06 [4.68E-06]	7.65E-06 [4.65E-06]	7.62E-06 [4.63E-06]	7.56E-06 [4.59E-06]
Model3	7.70E-06* [4.67E-06]	7.67E-06* [4.65E-06]	7.64E-06* [4.63E-06]	7.57E-06* [4.59E-06]
Model4	7.70E-06* [4.68E-06]	7.67E-06* [4.66E-06]	7.64E-06* [4.64E-06]	7.57E-06* [4.60E-06]
Model5	7.69E-06* [4.67E-06]	7.66E-06* [4.65E-06]	7.63E-06* [4.63E-06]	7.57E-06* [4.59E-06]
Model6	7.70E-06* [4.68E-06]	7.67E-06* [4.66E-06]	7.64E-06* [4.64E-06]	7.57E-06* [4.60E-06]
Model7	7.71E-06* [4.67E-06]	7.68E-06* [4.65E-06]	7.65E-06* [4.63E-06]	7.59E-06* [4.59E-06]
Model8	7.71E-06* [4.67E-06]	7.68E-06* [4.65E-06]	7.65E-06* [4.63E-06]	7.59E-06* [4.59E-06]

Note: Each cell contains the Clark and West statistics and the corresponding standard errors in square brackets, comparing Bad Realized Volatility and Good Realized Volatility, under the null that both forms realized volatilities do not differ markedly, one from the other. Positive and statistically significant estimates indicate out-performance of good realized volatility over the bad realized volatility. Also, * indicates statistical significance at 10% level.

Following from the foregoing, given that the MAT-HAR model variants have been shown to consistently outperform the conventional HAR model and trivially, the T-HAR model variants that assume threshold variables that are constant over time, we proceed to ascertain which of the model variants is optimal in the prediction of realized volatility, good realized volatility and bad realized volatility. The optimal model is chosen based on the Akaike information criterion, which selects *Model6* for the case of realized volatility, *Model4* in the case of good realized volatility and *Model1* (the conventional HAR model) in the case of bad realized volatility. The result is presented in Table 7, showing the estimated coefficients and the corresponding t-statistics. In the MAT-HAR predictive model for realized volatility, we find existence of economically and statistically significant threshold effect in both monthly and yearly frequencies, given the rejection of the null hypothesis of no threshold effect. There seems to be not threshold effect in the quarterly sampling frequency, when MAT-HAR model is used to predict realized volatility. In the case of the optimal MAT-HAR model for predicting good realized volatility, Model4 was preferred, indicating the existence of threshold effect at the year sampling frequency only. However, for the bad realized volatility, no MAT-HAR model variant is found to outperform the conventional HAR model, which is also indicative that no threshold effect exists for the bad realized volatility. Overall, whenever thresholds exist, it would be more appropriately captured using the moving average framework rather than using assuming a constant valued threshold or worse still, ignoring accounting for threshold effect completely.

Table 7: Parameter Estimates of AIC-based Optimal Model

Parameters	**Realized Volatility**	**Good Realized Volatility**	**Bad Realized Volatility**
Optimal Model	*Model6*	*Model4*	*Model1*
Lag Length	37	37	37
Number of Observations	1414	1414	1414
$\beta^{(0)}$	0.0008[3.6022]	0.0004[4.1141]	0.0005[4.2472]
Base Parameters			
$\beta^{(1)}$	0.7948[2.5240]	0.5641[3.9632]	0.3512[1.8486]
$\beta^{(2)}$	0.0511[0.3627]	0.0120[0.0909]	0.1179[0.9602]
$\beta^{(3)}$	-0.1420[-1.3332]	0.0449[0.5191]	0.1915[2.4936]
Threshold Parameters			
$\psi^{(1)}$	-0.2643[-1.4542]	-	-
$\psi^{(2)}$	-	-	-
$\psi^{(3)}$	0.2951[4.0394]	0.1681[2.1115]	-
R-Square	0.4003	0.4450	0.2377
Wald p-value	0.0003	0.0347	-

Note: The figures in square bracket are t-statistics, which are computed using Newey and West (1987) HAC estimator combined with Newey and West (1994) automatic bandwidth selection. Contending models are compared with the conventional HAR model.

5. Conclusion

This study set out to examine the forecast performance of a novel moving average threshold heteroscedastic autoregressive (MAT-HAR) model, in the prediction of the realized volatility of the US stock market. The models are fitted with three different sampling frequencies (monthly $(m=1)$, quarterly $(m=3)$ and yearly $(m=12)$), as consistent with the HAR model framework. This novel approach adopts moving averages, rather than values that are constant over time, to capture inherent threshold effects; as the latter are perceived to be too restrictive and unrealistic, when it comes to predicting the dynamics of realized volatility of the US stock market. In a bid to ascertain the superiority of the novel MAT-HAR model variants over the conventional HAR model and the T-HAR model variants, we subject all the models to predictability performance test, using the conventional RMSE for individual forecast error evaluation and the Clark and West (2007) test statistic for a pairwise comparison, given that the model being examined are nested. Consequently, the unrestricted model here is the novel MAT-HAR model, while the conventional HAR model and the T-HAR model variants are presumed to be the restricted versions, as the case may be. The variant of these (time-varying or time in-variant) threshold HAR models differ, one from the other, on the bases of the sampling frequency for which thresholds are incorporated (either singly, in pairs or all three at a time).

We attempt to answer two pertinent questions, hinging on model performance or outperformance. The first focuses on the relevance, or otherwise, of accounting for threshold effects when predicting the realized volatility of the US stock market. Consequently, we inquire, "Does threshold matter in the HAR-based Realized volatility of stock market?" and compare the different threshold HAR models the T-HAR variants and the MAT-HAR variants separately with the conventional HAR model (our baseline model). The intuition is to see if accounting for thresholds, if and whenever they exist, at specified sampling frequencies, does improve the predictive performance of the threshold HAR models over the conventional HAR model. The second borders on the relevance of accounting for asymmetry in realized volatility. "Does asymmetry matter in the MAT-HAR?" We subsequently estimate, in addition to the conventional HAR model (*Model1*), seven different models *(Model2 – Model8)* each for MAT-HAR and T-HAR model constructs, such that *Model2, Model3* and *Model4* specify the thresholds (constant or time varying) at sampling frequencies 1, 2 and 3, respectively, while *Model5, Model6* and *Model7* specified the thresholds in pairs of sampling frequencies corresponding to sampling frequencies 1&2, 1 & 3 and 2 & 3, respectively, and *Model8* assumes threshold effect exists at the three sampling frequencies. The forecast errors of all eight models using the RMSE and Clark and West (2007) statistic in the in-sample and out-of-sample periods were examined for all measures of realized volatility.

We find some interesting results, which do not differ too markedly from our expectations. First, a HAR model that incorporates a time invariant threshold variable is less likely to outperform the conventional HAR model that does not account for any threshold parameters. Thus, accounting for threshold effects using constant threshold values may not matter. Second, time-varying thresholds are more likely to better predict the realized volatility than the conventional and trivial, T-HAR variants. Third, asymmetry does matter, as good realized volatility and bad realized volatility cannot be predicted in the same way. Fourth, the optimal MAT-HAR model for predicting the realized volatility of the US stock market is *Model6* (threshold in pair of sampling frequencies (1&3)) and Model4 (threshold in a specific sampling frequency (3)) for realized volatility and good realized volatility, respectively. Finally, the outperformance stance of the MAT-HAR model variants is not sensitive to the choice of sample periods and realized volatility measures.

It must be realized that appropriate modeling and accurate forecasting of volatility is of importance due to several reasons, as pointed out by Poon and Granger (2003): (1) when volatility is interpreted as uncertainty, it becomes a key input to investment decisions and portfolio choices; (ii) volatility is the most important variable in the pricing of derivative securities, since to price an option, one needs reliable estimates of the volatility of the underlying assets; (iii) financial risk management according to the Basle Accord of 1996 also requires modeling and forecasting of volatility as a compulsory input to risk-management for financial institutions around the world, (iv) financial market volatility, can have widespread repercussions on the economy as a whole, as witnessed during the recent “Great Recession”, and the “Great Depression” earlier, via its effect on real economic activity and public confidence. Hence, estimates of market volatility can serve as a measure for the vulnerability of financial markets and the economy, and can help policy makers design appropriate policies. Naturally, appropriate modeling and accurate forecasting of the process of volatility has ample implications for portfolio selection, the pricing of derivative securities and risk management, and as we show the MAT-HAR model can help in all these contexts.

As part of future analysis, it would be interesting to extend our study to other financial markets of both developed and emerging economies, as well as to commodity markets.

Data Availability Statement:

The data that support the findings of this study are openly available in the MeasuringWorth database at: https://www.measuringworth.com/datasets/DJA/index.php.

References

Alam, M. S., Shahzad, S. J. H., & Ferrer, R. (2019). Causal flows between oil and forex markets using high-frequency data: Asymmetries from good and bad volatility. Energy Economics, *104513.* doi:10.1016/j.eneco.2019.104513.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2001). The Distribution of Realized Exchange Rate Volatility. Journal of the American Statistical Association, 96(453), 42–55.

Apergis, N., Baruník, J., & Lau, M. C. K. (2017). Good volatility, bad volatility: What drives the asymmetric connectedness of Australian electricity markets? Energy Economics, 66, 108–115.

Baruník, J., Kočenda, E., & Vácha, L. (2016). Asymmetric connectedness on the U.S. stock market: Bad and good volatility spillovers. Journal of Financial Markets, 27, 55–78.

Bauer, G. H., & Vorkink, K. (2011). Forecasting multivariate realized stock market volatility. Journal of Econometrics, 160(1), 93–101.

BenSaïda, A. (2018). Good and bad volatility spillovers: An asymmetric connectedness. Journal of Financial Markets, 43, 78-95.

Campbell, J.Y. & Thompson, S.B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? Review of Financial Studies 21, 1509–1531.

Chen, Y., Ma, F., & Zhang, Y. (2019). Good, bad cojumps and volatility forecasting: New evidence from crude oil and the U.S. stock markets. Energy Economics, 81, 52-62.

Cho, S., & Shin, D. W. (2016). An integrated heteroscedastic autoregressive model for forecasting realized volatilities. Journal of the Korean Statistical Society, 45(3), 371–380.

Clark, T.E. & West, K.D. (2007). Approximately normal tests for equal predictiveaccuracy in nested models. Journal of Econometrics 138, 291–311.

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

Corsi, F., Audrino, F. and Reno, R. (2012). HAR modeling for realized volatility forecasting. In *Handbook of Volatility Models and Their Applications*, ed. By Bauwens, L., Hafner, C. M. and Laurent S., Chap. 15, pp.363-382. New Jersey, USA: John Wiley & Sons, Inc.

Degiannakis, S. (2017). The one-trading-day-ahead forecast errors of intra-day realized volatility. Research in International Business and Finance, 42, 1298–1314.

Degiannakis, S., & Filis, G. (2017). Forecasting oil price realized volatility using information channels from other asset classes. Journal of International Money and Finance, 76, 28–49.

Degiannakis, S. (2018). Multiple days ahead realized volatility forecasting: Single, combined and average forecasts. Global Finance Journal, 36, 41–61.

Gkillas, K., Gupta, R., & Pierdzioch, C. (2019a). Forecasting realized gold volatility: Is there a role of geopolitical risks? Finance Research Letters, in press, doi.org/10.1016/j.frl.2019.08.028

Gkillas, K., Gupta, R., & Pierdzioch, C. (2019b). Forecasting (downside and upside) realized exchange-rate volatility: Is there a role for realized skewness and kurtosis? Physica A: Statistical Mechanics and Its Applications, 532, 121867.
Gong, X., & Lin, B. (2017). Forecasting the good and bad uncertainties of crude oil prices using a HAR framework. Energy Economics, 67, 315–327.
Gong, X., & Lin, B. (2019). Modeling stock market volatility using new HAR-type models. Physica A: Statistical Mechanics and Its Applications, 516, 194-211.
Gupta, R., Pierdzioch, C., Selmi, R., & Wohar, M. E. (2018). Does partisan conflict predict a reduction in US stock market (realized) volatility? Evidence from a quantile-on-quantile regression model. The North American Journal of Economics and Finance, 43, 87–96.
He, F., Wang, Z. & Yin, L. (2019). Asymmetric volatility spillovers between international economic policy uncertainty and the U.S. stock market. The North American Journal of Economics and Finance, In Press, doi.org/10.1016/j.najef.2019.101084.
Ma, F., Zhang, Y., Huang, D., & Lai, X. (2018). Forecasting oil futures price volatility: New evidence from realized range-based volatility. Energy Economics, 75, 400-409.
Ma, F., Wahab, M. I. M., Huang, D., & Xu, W. (2017). Forecasting the realized volatility of the oil futures market: A regime switching approach. Energy Economics, 67, 136–145.
Mei, D., Liu, J., Ma, F., & Chen, W. (2017). Forecasting stock market volatility: Do realized skewness and kurtosis help? Physica A: Statistical Mechanics and Its Applications, 481, 153–159.
Motegi, K., Cai, X., Hamori, S. and Xu, H. (2019). Moving Average Threshold Heterogeneous Autoregressive (MAT-HAR) Models. Available at http://dx.doi.org/10.2139/ssrn.3426182
Newey, W. K., & West, K. D. (1987): A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica, 55, 703-708.
Newey, W. K., & West, K. D. (1994): Automatic Lag Selection in Covariance Matrix Estimation. Review of Economic Studies, 61, 631-653.
Oloko, T.F. (2018). Portfolio diversification between developed and developing stock markets: The case study of US and UK investors in Nigeria. Research in International Business and Finance, 45, 219-232.
Peng, H., Chen, R., Mei, D., & Diao, X. (2018). Forecasting the realized volatility of the Chinese stock market: Do the G7 stock markets help? Physica A: Statistical Mechanics and Its Applications, 501, 78–85.
Ping, Y., & Li, R. (2018). Forecasting realized volatility based on the truncated two-scales realized volatility estimator (TTSRV): Evidence from China's stock market. Finance Research Letters, 25, 222–229.
Poon, S-H, & Granger, C. W. J. (2003). Forecasting Volatility in Financial Markets: A Review. Journal of Economic Literature, 41(2), 478-539.
Qu, H., Duan, Q., & Niu, M. (2018). Modeling the volatility of realized volatility to improve volatility forecasts in electricity markets. Energy Economics, 74, 767–776.
Salisu, A. A. & Isah, K. O. (2017). Revisiting the oil price and stock market nexus: A nonlinear Panel ARDL approach. Economic Modelling, 66, 258–271.

Salisu, A. A. & Ogbonna, A. E. (2018). Does time-variation matter in the stochastic volatility components for G7 stock returns. *Working Papers 062, Centre for Econometric and Allied Research, University of Ibadan*

Salisu, A. A. & Ogbonna, A. E. (2019). Another look at the energy-growth nexus: New insights from MIDAS regressions. Energy, 174, 69-84.

Salisu, A. A. & Oloko, T.F. (2015). Modeling oil price-US stock nexus: A VARMA-BEKK-AGARCH approach. Energy Economics, 50, 1-12.

Salisu, A. A., Swaray, R. and Oloko, T. F. (2019). Improving the predictability of the oil–US stock nexus: The role of macroeconomic variables. Economic Modelling, 76, 153-171.

Salisu A. A., Adediran, I., A., Oloko, T. and Ohemeng, W. (2019). The heterogeneous behaviour of the inflation hedging property of cocoa. North American Journal of Economic and Finance, in Press, https://doi.org/10.1016/j.najef.2019.101093

Sharma, P., & Vipul. (2016). Forecasting stock market volatility using Realized GARCH model: International evidence. The Quarterly Review of Economics and Finance, 59, 222–230.

Solnik, B.K. (1974). Why not diversify international rather than domestically? Financial Analysis Journal, 89–94.

Tian, F., Yang, K., & Chen, L. (2017). Realized volatility forecasting of agricultural commodity futures using the HAR model with time-varying sparsity. International Journal of Forecasting, 33(1), 132–152.

Tong, H. (1978). On a threshold model. In Pattern Recognition and Signal Processing, ed. By Chen, C. H. Sijthoff and Noordhoff, Amsterdam.

Wu, X., & Hou, X. (2019). Forecasting realized variance using asymmetric HAR model with time-varying coefficients. Finance Research Letters, 30, 89-95.

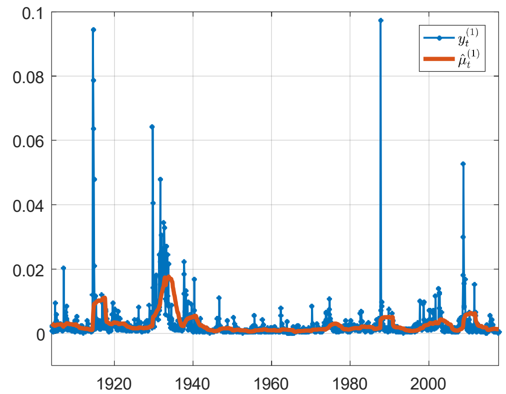
Wu, X., & Wang, X. (2019). Forecasting volatility using realized stochastic volatility model with time-varying leverage effect. Finance Research Letter, In Press, doi.org/10.1016/j.frl.2019.08.019

Yao, X., Izzeldin, M., & Li, Z. (2019). A novel cluster HAR-type model for forecasting realized volatility. International Journal of Forecasting, 35(4), 1318–1331.

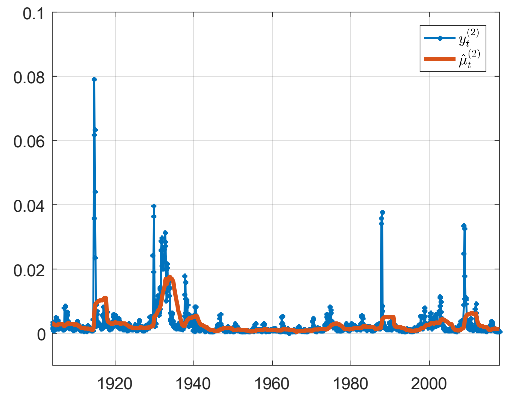
Zhou, W., Pan, J., & Wu, X. (2019). Forecasting the realized volatility of CSI 300. Physica A: Statistical Mechanics and Its Applications, 531, 121799.

[Dataset] MeasuringWorth; 2019: https://www.measuringworth.com/datasets/DJA/index.php.

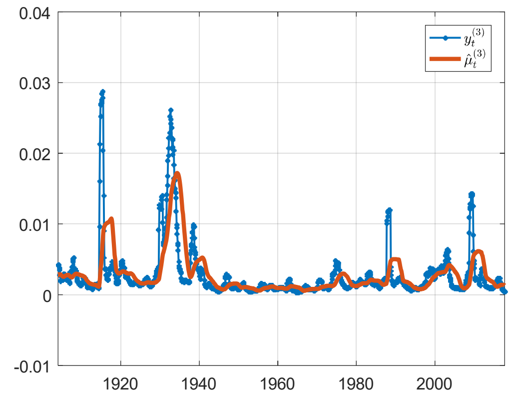
Appendix



(a) *Month* $(k=1)$

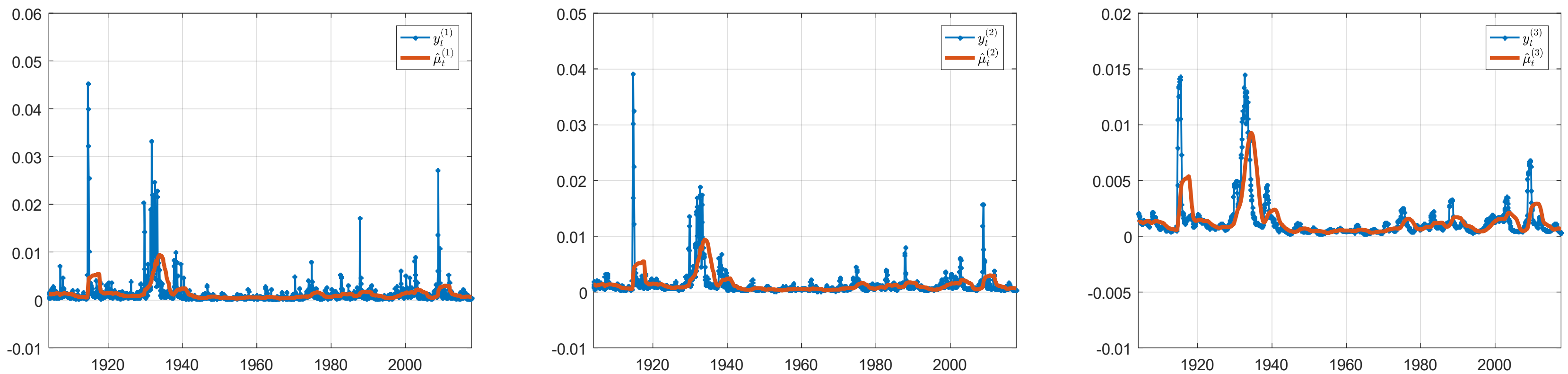


(b) *Quarter* $(k=2)$



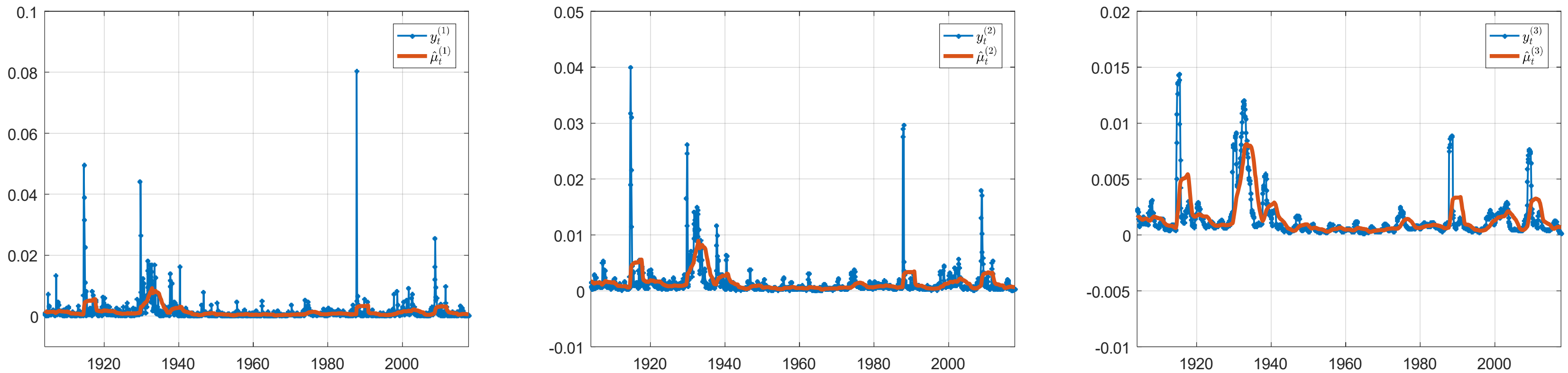
(c) *Year* $(k=3)$

Figure 1: Time Plot of Realized Volatility



(a) *Month* $(k=1)$ (b) *Quarter* $(k=2)$ (c) *Year* $(k=3)$

Figure 2: Time Plot of Good Realized Volatility



(a) *Month* $(k=1)$ (b) *Quarter* $(k=2)$ (c) *Year* $(k=3)$

Figure 3: Time Plot of Bad Realized Volatility