北京大学中国经济研究中心
China Center for Economic Research
讨论稿系列
Working Paper Series

No.E2015007 **October 2015**

Modeling Long Memory Volatility Using Realized Measures of Volatility: A Realized HAR GARCH Model

Zhuo Huang Hao Liu Tianyi Wang

Abstract

Long memory is an important feature of the volatility of financial returns. We document that the recently developed Realized GARCH model (Hansen et al. 2012) is insufficient for capturing the long memory of underlying volatility. We develop a parsimonious variant of the Realized GARCH model by introducing the HAR specification of Corsi (2009) into the volatility dynamics. A comparison of the theoretical and sample autocorrelation functions shows that the new model specification better captures the long memory dynamics of volatility. We calculate the multi-period out-of-sample volatility forecasts for several return series and find that the new model is a significant improvement over the classic Realized GARCH model.

Keywords: Realized GARCH, HAR, Long Memory, Realized Kernel

Modeling Long Memory Volatility Using Realized Measures of Volatility: A Realized HAR GARCH Model*

Zhuo Huang[a] Hao Liu[a] Tianyi Wang[b]

[a] *Peking University, National School of Development*

[b] *University of International Business and Economics*

School of Banking and Finance

Abstract

Long memory is an important feature of the volatility of financial returns. We document that the recently developed Realized GARCH model (Hansen et al. 2012) is insufficient for capturing the long memory of underlying volatility. We develop a parsimonious variant of the Realized GARCH model by introducing the HAR specification of Corsi (2009) into the volatility dynamics. A comparison of the theoretical and sample autocorrelation functions shows that the new model specification better captures the long memory dynamics of volatility. We calculate the multi-period out-of-sample volatility forecasts for several return series and find that the new model is a significant improvement over the classic Realized GARCH model.

Keywords: Realized GARCH, HAR, Long Memory, Realized Kernel

*Zhuo Huang acknowledges financial support from the Youth Fund of the National Natural Science Foundation of China (71201001) and from the Ministry of Education of China, Humanities and Social Sciences Youth Fund (12YJC790073). Tianyi Wang (corresponding author) acknowledges financial support from the Youth Fund of National Natural Science Foundation of China (71301027), the Ministry of Education of China, Humanities and Social Sciences Youth Fund (13YJC790146), the Fundamental Research Fund for the Central Universities in UIBE(14YQ05) and additional support from UIBE (XK2014116). Email: tianyiwang@uibe.edu.cn. We thank the editor and anonymous referees for their useful comments that greatly improve the quality of this paper.

1 Introduction

Accurately modeling and forecasting asset return volatility is a critical skill in asset pricing, risk management, asset allocation, and other financial applications. Since the seminal work by Engle (1982) and Bollerslev (1986),GARCH volatility models have been developed and extended to explain various stylized facts of financial return and volatility, such as volatility clustering, leverage effects, leptokurtosis, and long memory.

With the development of electronic trading, complete transaction and quote record data can be accessed easily, which has encouraged the use of high frequency data. Andersen and Bollerslev (1998) show that the daily aggregated squared intraday returns can be used as an accurate measure of latent volatility. During the past fifteen years, many studies have tried to exploit the superior information contained in a wide range of realized measures of volatility such as the 5-minute realized variance or realized kernel (RK, see Barndorff-Nielsen et al. (2008)) to model and forecast volatility. Engle (2002) introduces the GARCH-X model, which includes the realized variance as an exogenous variable in the volatility dynamics equation. To complete the GARCH-X model so that multi-step volatility prediction is feasible, Engle and Gallo (2006) develop a multiplicative error model by specifying a separate dynamic equation for the realized measure. Based on the same structure, Shephard and Sheppard (2010) propose another extension of the GARCH-X model, called the HEAVY model. Çelik and Ergin (2014) conduct a thorough comparison of the forecasting performances of various volatility models that use high frequency data.

Using a similar but more natural approach, Hansen et al. (2012) develop the Realized GARCH model, in which a measurement equation is introduced to relate the realized measure directly to the latent conditional volatility. The Realized GARCH model not only outperforms the other GARCH family models in terms of modeling and forecasting volatility, it also maintains many attractive features of the conventional GARCH framework. A number of studies have extended the classic Realized GARCH model and the model has been used in empirical finance research. Hansen et al. (2014) develop a bi-variate version of the Realized GARCH to model conditional variance and conditional Beta of stock returns. Hansen and Huang (2015) extend it to the Realized EGARCH model by adopting a more flexible leverage function and allowing multiple realized measures. Watanabe (2012) and Louzis et al. (2014) assume a skewed-t distribution in the Realized GARCH model and show its advantages for predicting Value-at-Risk. Lunde and Olesen (2013) apply the model to forecast volatility in electricity markets. Banulescu et al. (2015) use the extended Realized GARCH model to identify major volatility shocks during the recent financial crisis. Hansen et al. (2015) demonstrate the capacity of the Realized GARCH framework to explain the variance risk premium suggested in the CBOE VIX index.

Despite the popularity of the Realized GARCH model, this study documents that the specification of the classic

Realized GARCH model is not sufficient for capturing the long memory, which is an important feature of volatility. Long memory refers to the fact that the autocorrelation of the squared or absolute returns of financial assets, as a proxy for underlying volatilities, decay at a slow rate. Financial volatility is widely accepted as a long memory process (Ding et al. 1993; Bollerslev and Ole Mikkelsen 1996; Breidt et al. 1998; Yalama and Çelik 2013). Baillie et al. (1996) propose a fractionally integrated GARCH model (FIGARCH) to model long memory volatility. Ebens (1999) models realized volatility with a well-established univariate ARFIMA framework. Andersen et al. (2003) use the fractional difference operators in FIGARCH models to model the realized variance. A similar structure is discussed with a time-varying extension in Belkhouja and Boutahary (2011). Corsi (2009) proposes a heterogeneous autoregressive (HAR) model of realized variance by arguing that a long memory process can be approximated using a short-memory specification with a cascade structure of lags[1]. Todorova (2015) applies an extended HAR model, in which the volatility of volatility is time-varying, to study the realized volatility dynamics of metal futures.

As a remedy, in this study we develop an extended model, the Realized HAR GARCH, by specifying the heterogeneous autoregressive (HAR) terms of realized volatility as in Corsi (2009). An empirical analysis of 29 return series (two stock indices and 27 Dow Jones Industrial Average (DJIA) individual stocks) suggests that adding the HAR terms can improve the model's ability to capture the long memory feature in both latent and realized volatility. Under several commonly used loss functions, we show that our new model specification outperforms the classic Realized GARCH (1,1) model in multi-period volatility forecasting, both in-sample and out-of-sample[2].

2 Econometric Methodology

2.1 Realized GARCH

Let r_t be the return of an asset on date t, x_t be the realized measure, and $\mathcal{F}_t$ be the information set generated by $\{r_t, x_t, r_{t-1}, x_{t-1}, \ldots\}$. Let $\tilde{h}_t$ denote $\log h_t$ and $\tilde{x}_t$ denote $\log x_t$. The Realized GARCH model is given by

$$\begin{aligned}
r_t &= \mu + \sqrt{h_t} z_t, \qquad h_t = var(r_t|\mathcal{F}_{t-1}) \\
\tilde{h}_t &= \omega + \beta \tilde{h}_{t-1} + \gamma \tilde{x}_{t-1} \\
\tilde{x}_t &= \xi + \phi \tilde{h}_t + \tau(z_t) + u_t
\end{aligned}$$

[1]We thank the anonymous referee to point this out.

[2] We also compared the forecasting performance between our new model and FIGARCH model, a classic long-memory volatility model within the GARCH framework and found the Realized HAR GARCH model empirically superior. Results are presented in an unpublished supplementary appendix

where $z_t \sim N(0,1)$, $u_t \sim N(0,\sigma_u^2)$, with z_t and u_t being independent. The second equation is the GARCH equation where the lagged the realized measure replaces the lagged square return. The third equation is the measurement equation that takes the realized measure as a measurement of the latent volatility. $\tau(z_t) = \tau_1 z_t + \tau_2(z_t^2 - 1)$ is the leverage function, which captures the asymmetric dependence between return and volatility. One advantage of this measurement equation is that the model can deliver close-to-close volatility without pre-manipulation of the opening hour data-based realized measures.

2.2 Realized HAR GARCH

In this section we adopt a HAR specification for the realized measures in the GARCH equation. The Realized HAR GARCH model is then defined by extending the volatility dynamic equation:

$$\tilde{h}_t = \omega + \beta\tilde{h}_{t-1} + \gamma_d\tilde{x}_{t-1} + \tfrac{\gamma_w}{4}\textstyle\sum_{i=2}^{5}\tilde{x}_{t-i} + \tfrac{\gamma_m}{17}\textstyle\sum_{i=6}^{22}\tilde{x}_{t-i} \tag{1}$$

After simple reparameterization, equation 1 can be written as:

$$\tilde{h}_t = \omega + \beta\tilde{h}_{t-1} + \gamma_d^*\tilde{x}_{t-1} + \tfrac{\gamma_w^*}{5}\textstyle\sum_{i=1}^{5}\tilde{x}_{t-i} + \tfrac{\gamma_m^*}{22}\textstyle\sum_{i=1}^{22}\tilde{x}_{t-i} \tag{2}$$

where $\gamma_d^* = \gamma_d - \frac{\gamma_w}{4}$,$\gamma_w^* = 5\left(\frac{\gamma_w}{4} - \frac{\gamma_m}{17}\right)$, and $\gamma_m^* = \frac{22}{17}\gamma_m$. Following Corsi (2009), those components of realized measures are related to information from the last day, last week, and last month. Unlike the Realized GARCH model with multiple realized measures, our variant has multiple lags of one realized measure. Therefore, only one measurement equation is needed. The current framework can be easily extended to multiple realized measures cases and we leave this for further studies. Compared to classic autoregressive specifications with 22 lags, the HAR structure has far fewer parameters, yet still explicitly preserves information from the 22 lags.

In contrast to the HAR model, the proposed model focuses on the latent volatility corresponding to inter-day returns rather than on realized measures corresponding to intra-day returns. For most applications, such as pricing derivatives and calculating Value-at-Risk, it is important to take overnight returns into account. Such an adjustment is not part of the HAR model, whereas the Realized HAR GARCH model makes the adjustment automatically via the measurement equation.

The model can easily preserve an ARMA type structure by substituting x_t in the GARCH equation:

$$\tilde{h}_t = \mu_h + (\beta + \phi\gamma_d)h_{t-1} + \frac{\phi\gamma_w}{4}\sum_{i=2}^{5}\tilde{h}_{t-i} + \frac{\phi\gamma_m}{17}\sum_{i=6}^{22}\tilde{h}_{t-i} + \gamma_d\epsilon_{t-1} + \frac{\gamma_w}{4}\sum_{i=2}^{5}\tilde{\epsilon}_{t-i} + \frac{\gamma_m}{17}\sum_{i=6}^{22}\tilde{\epsilon}_{t-i} \tag{3}$$

where $\mu_h = \omega + \xi(\gamma_d + \gamma_w + \gamma_m)$ and $\epsilon_t = \tau(z_t) + u_t$. This formula is a restricted ARMA(22,22) for $\tilde{h}_t$. Compared with its original Realized GARCH counterpart[3], the new volatility dynamics is more flexible in capturing the autocorrelation structure of the latent volatility.

For the measurement equation, we have

$$\tilde{x}_t = \mu_x + (\beta + \phi\gamma_d)\tilde{x}_{t-1} + \frac{\phi\gamma_w}{4}\sum_{i=2}^{5}\tilde{x}_{t-i} + \frac{\phi\gamma_m}{17}\sum_{i=6}^{22}\tilde{x}_{t-i} + \epsilon_t - \beta\epsilon_{t-1} \tag{4}$$

where $\mu_x = \phi\omega + \xi(1-\beta)$. Interestingly, the reduced form of the dynamics of $\tilde{x}_t$ is not AR(22) as the HAR model suggests. It is an ARMA (22,1) model with certain parameter constraints. This implies, although this is not the main concern of this study, that the Realized HAR GARCH model may generate richer dynamics for realized measures than the original HAR model.

3 Data

We use data on the daily close-to-close returns and daily realized measures of the two indices (S&P 500 index and NASDAQ 100 index) and 27 liquidly traded individual stocks to illustrate the performance of our new model [4] We use the realized kernel (RK) proposed by Barndorff-Nielsen et al. (2008) as the realized measure because it is robust to market microstructure noise. Our sample period is from 2002/01 to 2013/12. As the first 22 observations are used to calculate monthly volatility, the effective sample covers 2002/02-2013/12.

Table 1 provides the summary statistics for the daily close-to-close returns and the logarithms of daily realized kernels (log(RK)). All of the returns exhibit excess kurtosis, indicating that the distribution is fat-tailed. One can easily extend our model with a fat-tailed distribution; however, we leave this for further research as the current study focuses on the correlation structure of volatility. For log(RK), the factional integrated parameter d is estimated with local Whittle estimator proposed by Robinson (1995).[5] In Table 2, we present the estimated values of d for the full sample and three sub-samples divided by the recent financial crisis period. Most of the series have $d > 0.5$, which indicates that log(RK) has infinite variance and appears to be a non-stationary long-memory process. In addition, we tend to find a larger d on average during the financial crisis period than in the other two sub-samples. Earlier volatility studies tend to find a long memory parameter d of the level or logarithm of realized variance slightly less than 0.5 and conclude that the realized variance is a stationary long memory process. However, in more recent

[3]The Realized GARCH's corresponding reduced form is a restricted ARMA(1,1).

[4] We use the same 27 individual stock as in Hansen and Huang (2015) and add two stock indices. The daily returns are collected from Yahoo Finance. The realized kernels of the two stock indices are collected from the Oxford-Man Institute's Realized Library, and the realized kernels of individual stocks are kindly provided by Professor Asger Lunde.

[5]Estimating d with the 2-step exact local Whittle estimator of Shimotsu (2010) gives very similar results.

Table 1: Summary statistics for return and log(RK)

		Return					Log(RK)					
	Obs	Mean	Std	Skew	Kurt	Med	Mean	Std	Q1	Med	Q3	d
AA	2976	-0.04	2.74	-0.19	8.16	0.00	1.14	0.86	0.55	1.00	1.55	0.66
AIG	2971	-0.07	4.34	-0.27	40.01	-0.03	1.07	1.29	0.15	0.88	1.77	0.66
AXP	2971	0.04	2.43	0.10	10.43	0.00	0.60	1.19	-0.32	0.37	1.35	0.71
BA	2971	0.05	1.88	0.07	3.69	0.07	0.60	0.83	-0.01	0.46	1.08	0.66
BAC	2980	-0.01	3.38	-0.34	23.52	0.03	0.70	1.30	-0.29	0.56	1.43	0.71
C	2979	-0.06	3.64	-0.50	34.12	0.00	0.90	1.26	-0.04	0.71	1.64	0.66
CAT	2975	0.05	2.10	-0.11	5.10	0.06	0.73	0.83	0.17	0.60	1.16	0.66
CVX	2975	0.05	1.67	0.02	13.63	0.11	0.31	0.84	-0.27	0.22	0.79	0.64
DD	2974	0.03	1.78	-0.24	5.85	0.03	0.50	0.85	-0.11	0.37	1.01	0.62
DIS	2972	0.05	1.91	0.28	6.11	0.05	0.54	0.89	-0.12	0.38	1.04	0.65
GE	2976	0.00	1.99	0.04	9.48	0.00	0.42	1.05	-0.37	0.24	1.01	0.70
HD	2973	0.02	1.90	0.04	5.99	0.01	0.61	0.88	-0.01	0.44	1.07	0.66
IBM	2972	0.02	1.52	0.02	6.64	0.02	0.11	0.85	-0.49	-0.05	0.55	0.66
INTC	2972	0.00	2.20	-0.55	7.51	0.00	0.84	0.80	0.29	0.73	1.28	0.64
JNJ	2973	0.03	1.16	-0.70	23.55	0.02	-0.27	0.88	-0.88	-0.43	0.19	0.68
JPM	2977	0.03	2.75	0.27	13.41	0.00	0.80	1.13	-0.03	0.58	1.45	0.67
KO	2972	0.03	1.24	0.07	11.23	0.04	-0.10	0.82	-0.66	-0.23	0.34	0.64
MCD	2973	0.05	1.46	-0.18	7.15	0.09	0.16	0.90	-0.49	0.07	0.72	0.65
MMM	2971	0.04	1.45	-0.22	5.57	0.06	0.12	0.80	-0.39	0.02	0.53	0.62
MRK	2970	0.02	1.74	-0.42	9.61	0.03	0.39	0.86	-0.21	0.28	0.87	0.61
MSFT	2971	0.02	1.80	0.11	8.28	0.00	0.45	0.81	-0.14	0.30	0.93	0.63
PG	2973	0.03	1.14	-0.21	6.77	0.03	-0.18	0.78	-0.71	-0.30	0.22	0.59
T	2967	0.02	1.63	0.25	7.43	0.06	0.37	0.95	-0.30	0.21	0.89	0.66
UTX	2971	0.05	1.62	0.19	5.64	0.05	0.31	0.81	-0.24	0.20	0.74	0.62
VZ	2972	0.02	1.58	0.12	6.94	0.04	0.31	0.91	-0.33	0.14	0.80	0.64
WMT	2973	0.02	1.31	0.16	5.35	0.02	0.07	0.82	-0.49	-0.06	0.49	0.65
XOM	2974	0.04	1.60	-0.02	12.63	0.07	0.24	0.87	-0.35	0.14	0.69	0.64
NASDAQ	2985	0.03	1.45	-0.04	4.74	0.09	-0.46	0.96	-1.14	-0.56	0.13	0.62
SP500	2982	0.02	1.29	-0.17	8.60	0.07	-0.50	1.06	-1.26	-0.61	0.11	0.61

Note: Returns are reported as percentage changes and RK is reported in the corresponding scale. Kurt is the excess kurtosis. Med is the median. Q1 and Q3 are the first and third quartiles, respectively. Obs is the effective sample size. The unbalanced sample size is caused by the trimming procedure that removes short trading days (where high-frequency data spanned less than 20,000 seconds). Due to the difference of trading activities, it is possible that different stocks have slightly different simple sizes. The difference between the two indices is due to the cleaning process described in Realized Library's website.

volatility studies, the empirical results on whether d is greater than 0.5 are mixed, depending on the asset classes, the sample periods and the sampling frequency of the realized measure. Empirical evidences of long memory and non-stationarity in volatility have also been reported by Yalama and Çelik (2013) and Baruník and Dvořáková (2015).

Table 2: Summary statistics of estimated fractional integrated parameter d for log(RK)

Sample	Mean	Std	Min	Q25%	Median	Q75%	Max
Full	0.650	0.029	0.588	0.628	0.651	0.663	0.715
Pre-crisis	0.657	0.048	0.524	0.650	0.666	0.681	0.726
Crisis	0.748	0.062	0.609	0.715	0.757	0.788	0.855
Post-crisis	0.592	0.053	0.455	0.573	0.597	0.627	0.680

Note: Full : 201202-201312, pre-crisis: 200202-200707, crisis:200708-200907, post-crisis: 200908-201312.

Figure 1 plots the return, realized kernel, and log(RK) for four series: the two indices, the stock with the highest average market cap (Exxon Mobil, XOM), and the stock with the highest average turnover (Alcoa Inc., AA) over the sampling period. The results for the remaining series are available upon request.

4 Estimation

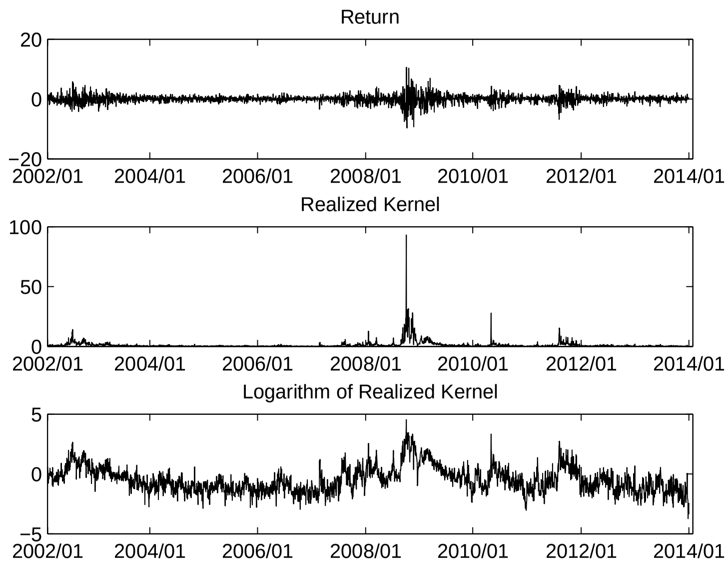
With the help of realized measures, the return and volatility shock (z_t, u_t) can be recursively determined. Given the initial value h_0, which we set as an additional parameter, we can calculate h_t by iteration. Therefore, the parameters can be estimated via the maximum likelihood estimation. The joint log-likelihood function is given by

$$\log L(\{r,x\}_{t=1}^{n};\theta) = \underbrace{-\frac{1}{2}\sum_{t=1}^{n}[\log(2\pi)+\log(h_t)+r_t^2/h_t]}_{=l(r)} \underbrace{-\frac{1}{2}\sum_{t=1}^{n}[\log(2\pi)+\log(\sigma_u^2)+u_t^2/\sigma_u^2]}_{=l(x|r)}.$$

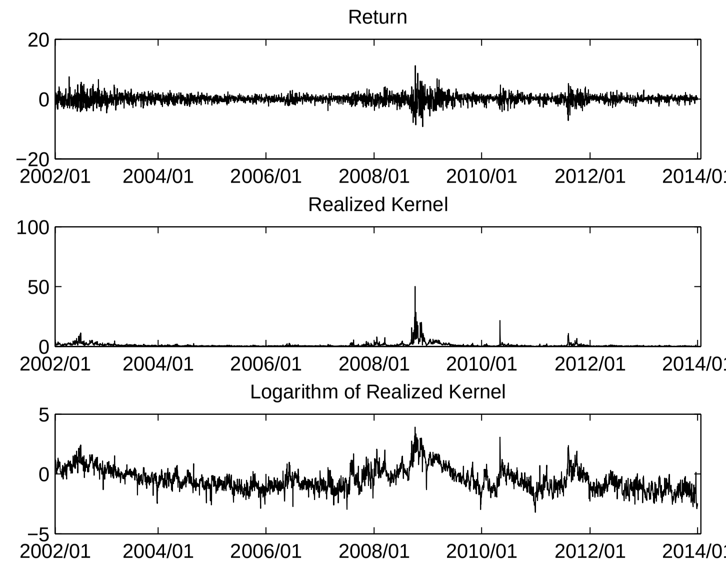
where the first part of the right hand-side of the equation is called partial likelihood since it only focuses on the goodness of fit for the returns.

As the estimated values of the long memory parameter d for log(RK) are mostly greater 0.5, the volatility equation in our model is augmented with a non-stationary long memory process. There has been no thorough investigation for the validity of the conventional MLE or QMLE statistical inference, such as the consistency and asymptotic distribution, for the Realized GARCH framework with non-stationary covariates.[6] Within the GARCH

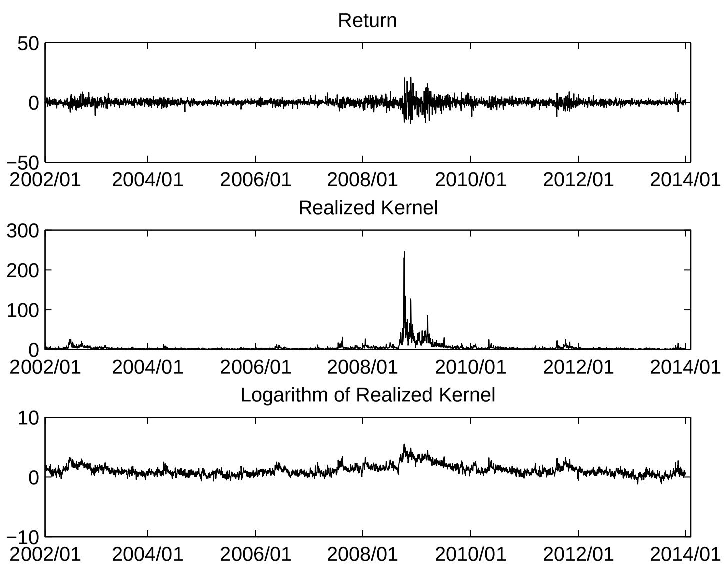
[6]We thank one of the referees for pointing out this.



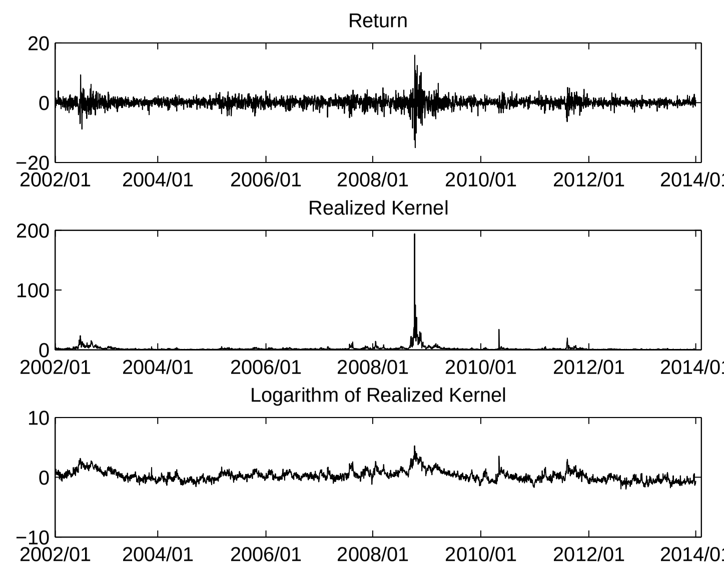
(a) S&P 500



(b) NASDAQ 100



(c) AA



(d) XOM

Figure 1: Return, realized kernel and log(RK) for different series

framework, Jensen and Rahbek (2004) show the conventional QMLE estimator for ARCH and GARCH models remains consistent with a Gaussian limit distribution regardless of the process being stationary or non-stationary. Han and Kristensen (2014) and Han (2015) investigate the asymptotic properties of GARCH-X models with possibly persistent and long memory covariates. They find that the standard inferential tools for the QMLE parameters are quite robust to the level of persistence of the covariates in the GARCH equation. In this paper we just follow Hansen et al. (2012) and calculate the conventional sandwich form robust covariance matrix and leave the thorough asymptotic analysis as an interesting question for further research.

4.1 Estimation Results

To save space, we only present the results for the S&P500 based using on close-to-close returns and log(RK) for the Realized HAR GARCH model in the full sample period:

$$\begin{aligned}
r_t &= \underset{(0.024)}{+0.023} + \sqrt{h_t} z_t \\
\tilde{h}_{t+1} &= \underset{(0.043)}{+0.252} + \underset{(0.093)}{0.388}\tilde{h}_t + \underset{(0.024)}{0.425}\tilde{x}_{t-1} + \underset{(0.073)}{0.114}\left(\frac{1}{4}\sum_{i=2}^{5}\tilde{x}_{t-i}\right) + \underset{(0.021)}{0.075}\left(\frac{1}{17}\sum_{i=6}^{22}\tilde{x}_{t-i}\right) \\
\tilde{x}_t &= \underset{(0.027)}{-0.417} + \underset{(0.032)}{0.953}\tilde{h}_t - \underset{(0.012)}{0.085} z_t + \underset{(0.009)}{0.116}(z_t^2 - 1) + u_t
\end{aligned}$$

with $\hat{\sigma}_u^2 = \underset{(0.007)}{0.237}$. The persistent parameter $\beta + \phi(\gamma_d + \gamma_w + \gamma_m) = 0.976$.

The corresponding estimates for the Realized GARCH model are:

$$\begin{aligned}
r_t &= \underset{(0.019)}{+0.024} + \sqrt{h_t} z_t \\
\tilde{h}_{t+1} &= \underset{(0.015)}{+0.161} + \underset{(0.024)}{0.601}\tilde{h}_t + \underset{(0.024)}{0.394}\tilde{x}_{t-1} \\
\tilde{x}_t &= \underset{(0.026)}{-0.417} + \underset{(0.031)}{0.950}\tilde{h}_t - \underset{(0.011)}{0.085} z_t + \underset{(0.009)}{0.117}(z_t^2 - 1) + u_t
\end{aligned}$$

with $\hat{\sigma}_u^2 = \underset{(0.019)}{0.240}$. The persistent parameter $\beta + \phi\gamma = 0.974$. The numbers in parentheses are the robust standard errors for each of the point estimates. Their signs and magnitudes are in accordance with the theoretical assumptions and the estimation result reported in Hansen et al. (2012)[7]. These calculations lead to several notable findings. 1)The Realized HAR GARCH model has significantly lower β when the persistent parameter increases a little. This means that better information contained in the lagged conditional variance can be partially explained by additional weekly and monthly components. 2) The statistical significance of these two terms can be established using the LR statistics

[7] Such as: 1) the persistent parameter is close to one, 2) the leverage effect exists indicated by a negative d_1. 3) β is much lower compared with transitional GARCH models that highlights the importance of realized measures.

[8] 3) $\gamma_d > \gamma_w > \gamma_m$ indicates a declining information content in the cascade structure. 4) The measurement equation remains almost the same across the different models.

Table 3: Full sample estimates for the Realized GARCH model

	μ	ω	β	γ_d	ξ	ϕ	d_1	d_2	σ_u^2
AA	0.001	0.214	0.565	0.410	-0.426	1.000	-0.054	0.063	0.143
AIG	-0.026	0.240	0.434	0.610	-0.324	0.884	-0.030	0.052	0.204
AXP	0.054	0.176	0.551	0.446	-0.369	0.978	-0.056	0.065	0.162
BA	0.071	0.174	0.612	0.353	-0.436	1.040	-0.046	0.079	0.147
BAC	0.008	0.239	0.416	0.613	-0.352	0.918	-0.060	0.048	0.173
C	-0.018	0.209	0.439	0.589	-0.305	0.914	-0.060	0.076	0.168
CAT	0.056	0.258	0.529	0.411	-0.549	1.079	-0.065	0.040	0.138
CVX	0.064	0.153	0.511	0.432	-0.315	1.063	-0.101	0.048	0.130
DD	0.035	0.117	0.532	0.461	-0.210	0.952	-0.070	0.047	0.152
DIS	0.054	0.145	0.602	0.381	-0.336	0.994	-0.071	0.051	0.156
GE	0.020	0.154	0.543	0.453	-0.312	0.969	-0.038	0.046	0.163
HD	0.045	0.123	0.571	0.429	-0.245	0.951	-0.043	0.051	0.147
IBM	0.027	0.179	0.523	0.467	-0.357	0.964	-0.062	0.034	0.140
INTC	0.019	0.168	0.515	0.510	-0.264	0.898	-0.034	0.032	0.130
JNJ	0.036	0.045	0.592	0.402	-0.122	0.961	-0.023	0.064	0.163
JPM	0.027	0.169	0.490	0.533	-0.280	0.924	-0.053	0.064	0.150
KO	0.030	0.070	0.556	0.443	-0.156	0.937	-0.047	0.061	0.153
MCD	0.060	0.099	0.661	0.311	-0.297	1.040	-0.059	0.075	0.168
MMM	0.042	0.151	0.549	0.400	-0.342	1.044	-0.065	0.040	0.156
MRK	0.025	0.192	0.598	0.347	-0.487	1.079	-0.053	0.051	0.181
MSFT	0.032	0.201	0.510	0.465	-0.377	0.985	-0.026	0.029	0.140
PG	0.028	0.062	0.537	0.392	-0.162	1.080	-0.047	0.058	0.160
T	0.037	0.080	0.547	0.445	-0.151	0.960	-0.061	0.063	0.177
UTX	0.048	0.137	0.550	0.441	-0.269	0.950	-0.052	0.069	0.156
VZ	0.030	0.078	0.589	0.387	-0.178	1.008	-0.054	0.061	0.164
WMT	0.015	0.079	0.615	0.331	-0.225	1.102	-0.027	0.059	0.145
XOM	0.048	0.131	0.507	0.449	-0.262	1.035	-0.104	0.050	0.130
NASDAQ	0.037	0.353	0.538	0.456	-0.758	0.950	-0.139	0.053	0.192
SP500	0.024	0.161	0.601	0.394	-0.417	0.950	-0.085	0.117	0.240
Average	**0.032**	**0.157**	**0.544**	**0.440**	**-0.320**	**0.986**	**-0.058**	**0.057**	**0.160**

Note: Average is the average value of each column. γ_d is significant at the 1% level in most cases.

The estimates for all 29 series in the full sample period are presented in Tables 3 and 4. To save space, we do not report the associated standard errors. Most of those parameters are significant, except for the mean in the return equation. The parameters associated with the realized measures are significant at the 1% level in most cases. The findings for the S&P500 series are similar to those in the averaged parameters and in most individual cases. The significance of the weekly and monthly components are established based on the LR statistic with $H_0 : \gamma_w = \gamma_m =$

[8] $H_0 : \gamma_w = \gamma_m = 0$. The corresponding LR statistic is 32.91 ($p < 0.001$). The test statistic is illustrated in Table 5.

Table 4: Full sample estimates for the Realized HAR GARCH model

	μ	ω	β	γ_d	γ_w	γ_m	ξ	ϕ	d_1	d_2	σ_u^2
AA	-0.001	0.328	0.271	0.460	0.133	0.116	-0.408	0.992	-0.056	0.062	0.141
AIG	-0.025	0.341	0.147	0.622	0.143	0.156	-0.333	0.901	-0.028	0.055	0.197
AXP	0.054	0.264	0.317	0.486	0.104	0.094	-0.371	0.980	-0.056	0.064	0.159
BA	0.070	0.332	0.234	0.416	0.180	0.115	-0.432	1.039	-0.047	0.078	0.144
BAC	0.009	0.334	0.134	0.647	0.158	0.121	-0.341	0.916	-0.058	0.046	0.168
C	-0.017	0.256	0.246	0.615	0.076	0.118	-0.290	0.910	-0.059	0.073	0.165
CAT	0.058	0.452	0.144	0.467	0.195	0.105	-0.545	1.077	-0.064	0.040	0.134
CVX	0.064	0.287	0.047	0.497	0.307	0.059	-0.304	1.054	-0.102	0.047	0.126
DD	0.035	0.194	0.178	0.516	0.221	0.092	-0.207	0.952	-0.070	0.045	0.149
DIS	0.053	0.319	0.083	0.468	0.285	0.142	-0.329	0.990	-0.072	0.050	0.152
GE	0.019	0.268	0.156	0.521	0.210	0.129	-0.297	0.958	-0.040	0.050	0.159
HD	0.045	0.219	0.204	0.491	0.211	0.108	-0.245	0.952	-0.043	0.052	0.144
IBM	0.027	0.293	0.214	0.512	0.201	0.068	-0.358	0.965	-0.062	0.034	0.138
INTC	0.016	0.284	0.113	0.574	0.234	0.145	-0.263	0.901	-0.036	0.029	0.126
JNJ	0.035	0.090	0.236	0.469	0.202	0.098	-0.124	0.956	-0.023	0.064	0.160
JPM	0.025	0.272	0.129	0.597	0.212	0.116	-0.275	0.923	-0.053	0.062	0.145
KO	0.030	0.126	0.203	0.500	0.199	0.123	-0.153	0.931	-0.047	0.061	0.150
MCD	0.062	0.212	0.249	0.393	0.167	0.145	-0.290	1.032	-0.058	0.075	0.165
MMM	0.043	0.267	0.191	0.459	0.184	0.098	-0.339	1.038	-0.064	0.040	0.152
MRK	0.026	0.367	0.203	0.410	0.178	0.121	-0.478	1.075	-0.052	0.051	0.177
MSFT	0.033	0.323	0.164	0.521	0.188	0.119	-0.358	0.968	-0.028	0.030	0.137
PG	0.027	0.116	0.161	0.448	0.207	0.087	-0.159	1.066	-0.048	0.058	0.157
T	0.040	0.131	0.173	0.512	0.192	0.132	-0.143	0.957	-0.059	0.062	0.172
UTX	0.047	0.224	0.228	0.494	0.173	0.113	-0.263	0.946	-0.053	0.068	0.153
VZ	0.031	0.134	0.270	0.442	0.161	0.099	-0.176	1.006	-0.054	0.061	0.162
WMT	0.016	0.157	0.222	0.391	0.193	0.100	-0.221	1.098	-0.027	0.058	0.142
XOM	0.048	0.228	0.119	0.501	0.259	0.056	-0.259	1.032	-0.105	0.049	0.127
NASDAQ	0.036	0.610	0.214	0.509	0.170	0.124	-0.753	0.945	-0.142	0.051	0.187
SP500	0.023	0.252	0.388	0.425	0.114	0.075	-0.417	0.953	-0.085	0.116	0.237
Average	**0.032**	**0.253**	**0.187**	**0.497**	**0.192**	**0.110**	**-0.295**	**0.986**	**-0.054**	**0.054**	**0.152**

Note: Average is the average value of each column. γ_d, γ_w, γ_m are significant at 1% level in most cases.

0. If we label the Realized HAR GARCH model as RHG and the Realized GARCH model as RG, then the statistic can be written as

$$LR = 2(\log L_{RHG} - \log L_{RG})$$

The statistic follows $\chi^2(2)$, under the null hypothesis. The results are reported in Table 5 for all of the series under different sample periods. Each column gives the total number of series with an LR statistic that is significant at different levels. For example, in the sub-prime crisis period, 15 out of 29 series reject the null at the 1% significance level; 8 of the remaining 14 series reject the null at the 5% significance level; 2 out of the remaining 6 series reject the null at the 10% significance level; and 4 series fail to reject the null even at the 10% level. Although the significance levels of the differences are lower during the sub-prime crises, the results highlight the importance of γ_w and γ_m in terms of model fit [9].

Table 5: Summary of the LR Test of $H_0 : \gamma_w = \gamma_m = 0$

	RHG***	RHG**	RHG*	RHG	Total
Full	29	0	0	0	29
Pre-crisis	29	0	0	0	29
Crisis	15	8	2	4	29
Post-crisis	29	0	0	0	29

Note: Full: 2002/02-2013/12; pre-crisis: 2002/02-2007/07; crisis: 2007/08-2009/07; post-crisis: 2009/08-2013/12. The model name in each row indicates the model with the larger loglikelihood.The following asterisks indicate the significance level of the difference: *** = 1%, ** = 5%, and *=10%.

4.2 Autocorrelation Functions of Conditional Variance and Realized Variance

In this section, we discuss the autocorrelation structure of the Realized HAR GARCH using plots of the autocorrelation function (ACF). As in Section 3, we only report the results for four selected series, i.e., S&P 500, NASDAQ 100, AA, and XOM. The results of the remaining series are very similar and available upon request. Figure 2 provides insight into whether the model is consistent with the correlation structure of its fitted latent volatility. The dash line is the ACF of the logarithmic latent volatility ($\tilde{h}_t$), calculated with the help of the population spectral density for ARMA models[10], as in equation 3 (and its Realized GARCH counterpart); the solid line is the sample-ACF (SACF) of the actual fitted $\tilde{h}_t$. If the model is internally consistent, the two lines should be close to each other. It is clear that the Realized GARCH is only able to capture the autocorrelation structure over the short-term, whereas the Realized

[9]The measurement equation of the Realized HAR GARCH model has slightly lower residual variances than the standard Realized GARCH. This probably indicates that the new specification also improves the fitting of the measurement equation. We thank the anonymous referee to point this out.

[10]See, for example, Section 6.1 of Hamilton (1994)

HAR GARCH can capture the structure over a much longer horizon.

Figure 3 provides insight into whether the model is able to capture the correlation structure of the market volatility. To do this, we plot the SACF of the logarithmic realized kernel $\tilde{x}_t$ with a solid line and the ACF of $\tilde{x}_t$ calculated using equation 4 and its Realized GARCH counterpart. It is obvious that for the classic Realized GARCH (1,1) model, the theoretical ACF starts to diverge from the SACF at around 10 lags[11]. In contrast, the Realized HAR GARCH model, which only introduces two more parameters, is able to capture the autocorrelation structure of the realized kernel and conditional variance for substantially longer horizons.

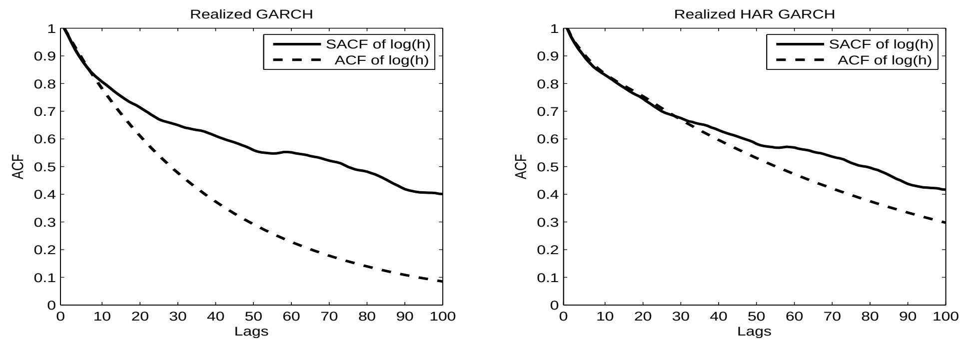
5 Multi-period Volatility Forecasting

Unlike the GARCH-X model, which can only produce a one-step ahead prediction of conditional variance, the more comprehensive Realized GARCH class models can predict multi-period conditional variance. Having demonstrated the superior performance of the Realized HAR GARCH model over the classic Realized GARCH model in fitting correlation structures, we now compare their out-of-sample volatility forecast performance. In this section, "Multi-period" means forecast conditional volatility in k days. In contrast to a one-step forecast, which can be calculated directly from equation 1, the log-linear specification complicates the multi-period forecast. Because $\mathbb{E}(\exp(\log h_{t+k})) \neq \exp \mathbb{E}(\log h_{t+k})$, the distributional aspect of $\log h_{t+k}$ must be considered to produce an unbiased forecast of h_{t+k}. To avoid the complexity of analytical formulas, we use a simple simulation method with empirical distributions to calculate the forecast. The last 520 observations are used for out-of-sample comparison, which means that each day in the out-of-sample period has 20 "future" volatilities to be compared. The procedure for obtaining $\hat{h}_{t+k}$ at time t is as follows:

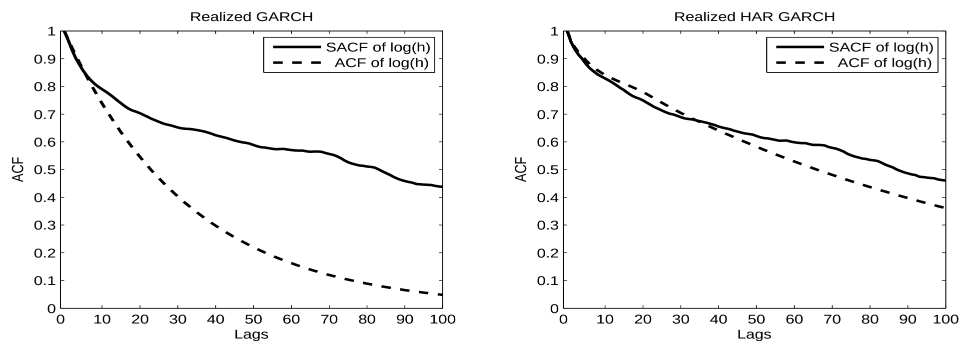
1. Estimate parameters Θ_t using estimation data ranging from $t-2399$ to t. The one-step forecast can be directly calculated by equation 1 and its Realized GARCH counterpart.

2. If $k > 1$, suppose we have the m-period forecast $\hat{h}_{t+m}$. Then calculate the 2400 pairs of $W_t = \{(z_s, u_s)|s = t, t-1, ..., t-2399\}$ based on the estimation data and Θ_t. Draw 5000 random pairs from W_t and simulate $\tilde{x}_{t+m}$ via the measurement equation (for the Realized HAR GARCH model, update the weekly volatility $\sum_{i=2}^{5} \tilde{x}_{t+m-i}$ and monthly volatility $\sum_{i=6}^{22} \tilde{x}_{t+m-i}$ accordingly). Calculate the logarithm conditional volatil-

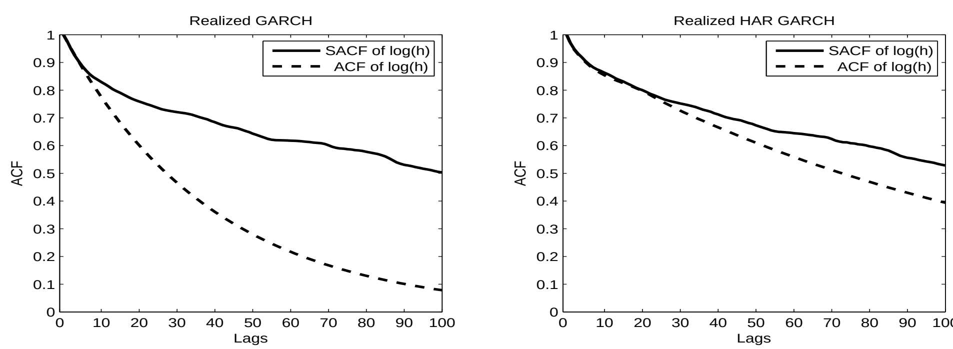
[11] We also check the Realized GARCH (p,q) with p, q up to 5, which has eight more parameters, but it fails to substantially alleviate the problem.



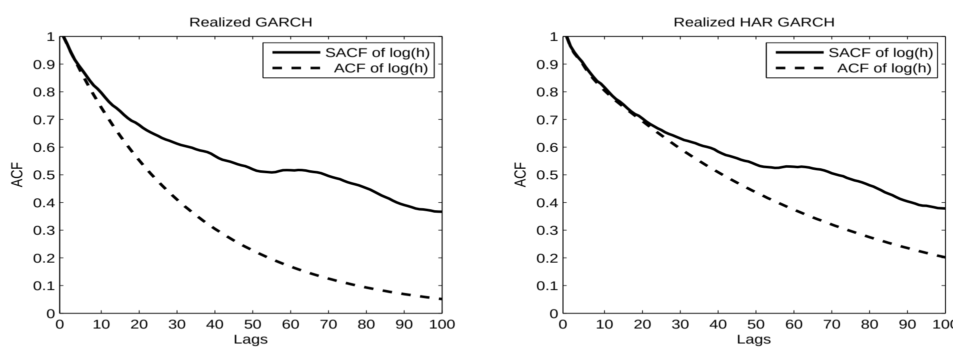
(a) S&P 500



(b) NASDAQ 100



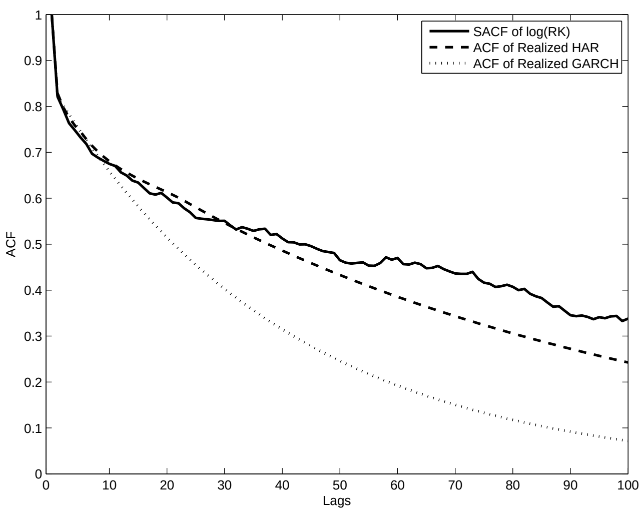
(c) AA



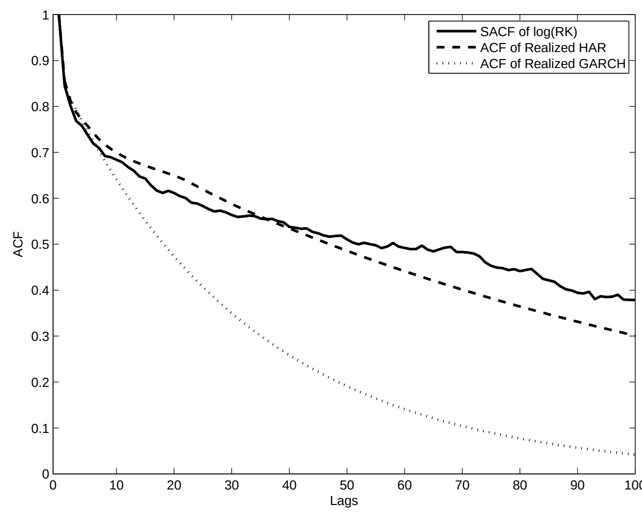
(d) XOM

Note: The dashed line is the theoretical ACF of $\log(h_t)$ and the solid line is the SACF of the actual fitted $\log(h_t)$. A large deviation between the two indicates an internal inconsistency in approximating the long-memory property.

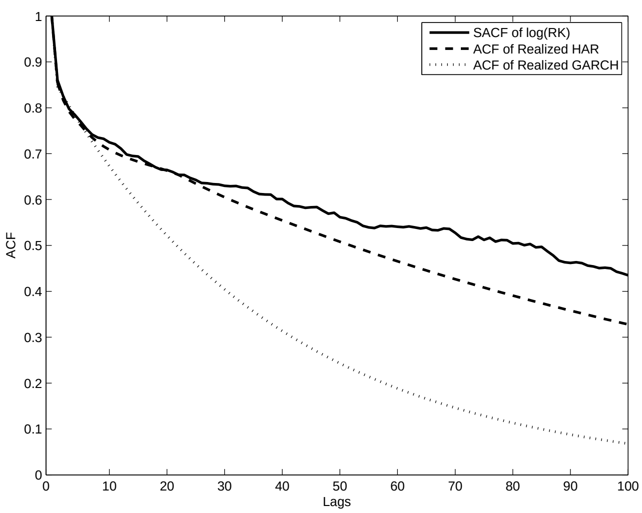
Figure 2: Theoretical and actual autocorrelation function on $\log(h_t)$ for the Realized GARCH and Realized HAR GARCH model



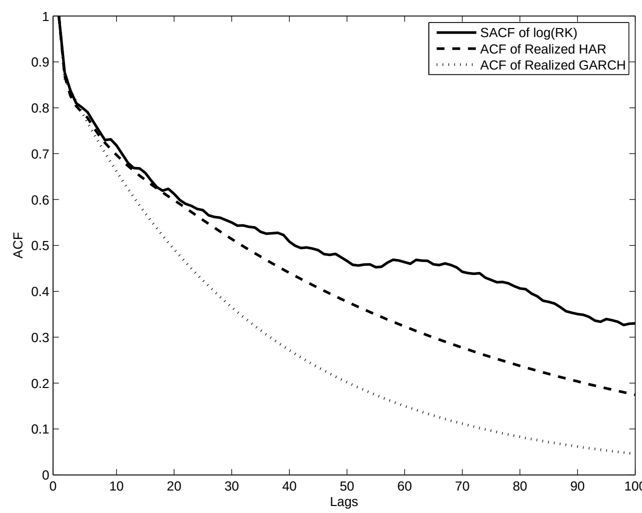
(a) S&P 500



(b) NASDAQ 100



(c) AA



(d) XOM

Note: The dashed line is the theoretical ACF of log(RK) for the Realized HAR GARCH model, the dotted line is the theoretical ACF of log(RK) for the Realized GARCH model and the solid line is the sample-ACF (SACF) of the actual log(RK). A large deviation indicating a lack of ability to replicate the correlation structure of market volatility.

Figure 3: Theoretical and actual autocorrelation function on log(RK) for the Realized GARCH and Realized HAR GARCH model

ity $\tilde{h}_{t+m+1}$. The $(m+1)$-period forecast $\hat{h}_{t+m+1}$ is then the sample average of 5000 simulated results:

$$\hat{h}_{t+m+1} = \frac{1}{5000} \sum_{1}^{5000} \exp(\tilde{h}_{t+m+1})$$

3. Repeat the process until $m = k$.

To evaluate the forecast performance, we use the adjusted realized kernel as the volatility proxy. The adjustment is needed because the realized kernel is calculated based on opening hour data, whereas the latent volatility measures the variation in close-to-close returns. In this study, we use a simple adjustment method by inflating the realized kernel to match the volatility of the close-to-close return on average. That is

$$RK_t^{Adj} = \frac{\sum_{s=1}^{T} r_s^2}{\sum_{s=1}^{T} RK_s} RK_t$$

where r_s and RK_s are the close-to-close returns and daily realized kernels, respectively.

We use three common evaluation criteria, mean squared error (MSE), mean absolute error (MAE), and the quasi-likelihood loss function (QLIKE), to compare the forecasting performance of the Realized GARCH model and the Realized HAR GARCH model. MSE and QLIKE are robust loss functions that provide consistent ranking with conditionally unbiased volatility proxies (Patton (2011)). The MAE, although popular, is not a robust loss function. The fact that QLIKE penalizes underestimation more than overestimation provides an additional motivation for using it to evaluate volatility models, especially for risk management applications. The definitions of the loss functions are as follows:

$$\begin{aligned}
MSE: \quad L(\hat{\sigma}^2, h) &= (\hat{\sigma}^2 - h)^2 \\
MAE: \quad L(\hat{\sigma}^2, h) &= |\hat{\sigma} - h| \\
QLIKE: \quad L(\hat{\sigma}^2, h) &= \log h + \frac{\hat{\sigma}^2}{h}
\end{aligned}$$

where $\hat{\sigma}$ is the volatility proxy which, in this study, is the adjusted realized kernel. h is the forecast for a given model. We compare their forecasting performances over four time horizons: one day, one week (5 days), two weeks (10 days), and one month (20 days). The multi-period forecasts are evaluated with the proxy and predicted conditional variance of the last day of the time horizon. As an example, the 10-day volatility forecast for all four series is provided in Figure 4.

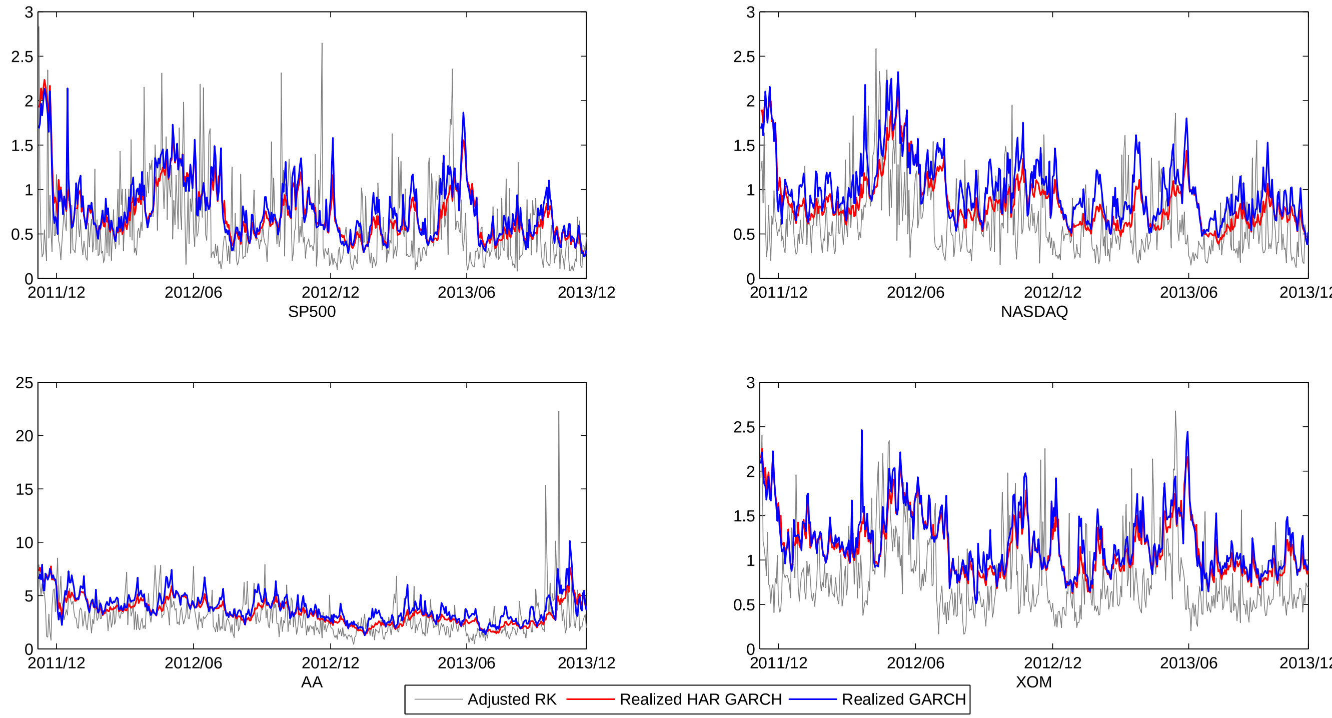


Figure 4: 10 days ahead volatility forecast for the Realized GARCH and Realized HAR GARCH model

It can be seen that the Realized GARCH model tends to report a higher volatility forecast than the Realized HAR GARCH model when the volatility is spiking up, as the Realized GARCH model can only use limited information about the correlation structure through γ_d. Therefore, the one-step forecast value becomes the dominating factor of the k-step forecast. The Realized HAR GARCH model performs forecasts using both the one-step forecast and the additional information about the correlation structure captured by γ_w and γ_m. These two parameters are trained by the data in the estimation window to react properly to volatility spikes in a relatively longer time horizon.

These visualized results are confirmed by all three loss functions. As seen in Table 6, the Realized HAR GARCH model outperforms the original Realized GARCH model for all four series under all forecast horizons. The statistical significance of the values of the loss functions is also calculated using Diebold-Mariano statistics. Each column gives the total number of series within the category indicated on the header line.

Table 6: Summary of k-steps ahead forecast difference via DM statistics

						Improvement			
	RHG***	RHG**	RHG*	RHG	RG	Mean	Q1	Median	Q3
RMSE									
k = 1	14	7	4	4	0	2.54%	1.41%	2.06%	3.70%
k = 5	20	3	3	3	0	6.77%	3.98%	6.08%	7.77%
k = 10	27	0	2	0	0	12.32%	7.26%	10.67%	13.37%
k = 20	29	0	0	0	0	19.07%	12.22%	19.07%	22.03%
MAE									
k = 1	22	6	0	1	0	3.73%	3.03%	3.87%	4.77%
k = 5	25	3	0	1	0	7.78%	5.59%	7.13%	9.35%
k = 10	29	0	0	0	0	14.99%	10.43%	14.77%	17.96%
k = 20	29	0	0	0	0	24.56%	17.07%	24.28%	28.83%
QLK									
k = 1	14	0	4	11	0	3.54%	1.69%	3.28%	5.29%
k = 5	17	5	7	4	1	6.62%	4.62%	6.40%	8.37%
k = 10	26	0	0	3	0	13.40%	10.47%	12.37%	17.36%
k = 20	28	0	1	0	0	24.15%	16.62%	23.48%	28.89%

Note: The model name in each column indicates the model with the best forecast accuracy. The following asterisks indicate the significance level of the difference: *** = 1%, ** = 5% and *=10%. The three panels are associated with the three loss functions: MSE (reported as RMSE), MAE, and QLK. For example, the last row indicates that for the 20-steps ahead forecast, RHG out preforms RG 28/29 times at the 1% significance level and 1/29 times at the 10% significance level under the QLK loss function. Q1 and Q3 are the first and third quartiles of the corresponding loss function's value.

The Diebold-Mariano statistics focus on the difference of loss function $d_t = L(\hat{\sigma}^2, h_t^{RG}) - L(\hat{\sigma}^2, h_t^{RHG})$ by testing whether the mean of d equals zero. The statistic is defined as

$$S = \frac{\bar{d}}{\sqrt{\widehat{LRV}/T}}$$

where $\bar{d} = \frac{1}{T}\sum_{t=1}^{T} d_t$, $LRV = \gamma_0 + 2\sum_{j=1}^{H} K\left(\frac{j}{H}\right) cov(d_t, d_{t-j})$. $K(.)$ is the Bartlett kernel function used in the Newy-West estimation of LRV, and H is the optimal bandwidth parameter estimated by Andrews (1991).

We also calculate the scaled difference of the loss functions as

$$Improvement = \frac{L(\hat{\sigma}^2, h^{RG}) - L(\hat{\sigma}^2, h^{RHG})}{L(\hat{\sigma}^2, h^{RG})}$$

The larger the value, the greater the "improvement" in the Realized HAR GARCH model over the original Realized GARCH model.

This study identifies several advantages of the Realized HAR GARCH model. 1) The Realized HAR GARCH outperforms the Realized GARCH in almost all cases (forecast horizons and loss functions). 2) The improvement is significant for a large proportion of all cases. 3) The significance of the difference tends to grow as the forecast horizon increases, eventually becoming significant at the 1% level. 4) This finding is confirmed by an increasing average "improvement". 5) Only 1 out of $3 \times 4 \times 29$ cases report the Realized GARCH as better than the Realized HAR GARCH. However, the difference is neither large (the scaled difference is 0.65%) nor statistically significant.

6 Conclusions

The recently proposed Realized GARCH model has received considerable attention in the financial econometrics literature. However, we determine that the benchmark model specification is not capable of modeling the long memory of volatility. We propose a parsimonious remedy, the Realized HAR GARCH model, which incorporates the HAR structure of realized variance into the GARCH equation. Using a dataset of 29 return series over a 12-year period that encompasses the dot-com bubble and the financial crisis, we compare the empirical performance of the Realized HAR GARCH model with that of the original Realized GARCH model. The results show that the new model specification is better for capturing the long memory feature of underlying volatility and provides more accurate multi-period out-of-sample volatility forecasts over different forecast horizons.

We leave two interesting questions for further research. First, FIGARCH models have been used on daily returns to model the long memory of underlying volatility. It would be interesting to incorporate the realized measures of volatility into the FIGARCH framework to exploit the more accurate information. Second, in this study we focus on the long memory volatility of stock returns. It has been documented that long memory exists in the volatilities of other asset classes such as bonds, exchange rates and commodities. Researchers could examine the empirical performance of our Realized HAR GARCH models in these markets.

References

Andersen, T. G. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review 39*, 885–905.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*(2), 579–625.

Andrews, D. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica 59*, 817–858.

Baillie, R. T., T. Bollerslev, and H. O. Mikkelsen (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 74*(1), 3–30.

Banulescu, G.-D., P. R. Hansen, Z. Huang, and M. Matei (2015). Volatility during the financial crisis through the lens of high frequency data: A realized garch approach. *Working Paper*.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica 76*(6), 1481–1536.

Baruník, J. and S. Dvořáková (2015). An empirical model of fractionally cointegrated daily high and low stock market prices. *Economic Modelling 34*, 193–206.

Belkhouja, M. and M. Boutahary (2011). Modeling volatility with time-varying figarch models. *Economic Modelling 28*, 1106–1116.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*(3), 307–327.

Bollerslev, T. and H. O. Mikkelsen (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics 73*, 151–184.

Bollerslev, T. and H. Ole Mikkelsen (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics 73*(1), 151–184.

Breidt, F. J., N. Crato, and P. De Lima (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics 83*(1), 325–348.

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

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

Ding, Z., C. W. Granger, and R. F. Engle (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance 1*(1), 83–106.

Ebens, H. (1999). Realized stock volatility. Technical report, The Johns Hopkins University, Department of Economics.

Engle, R. (2002). New frontiers for arch models. *Journal of Applied Econometrics 17*(5), 425–446.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica 50*, 987–1007.

Engle, R. F. and G. M. Gallo (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics 131*(1-2), 3–27.

Hamilton, J. D. (1994). *Time series analysis*, Volume 2. Princeton university press Princeton.

Han, H. (2015). Asymptotic properties of garch-x processes. *Journal of Financial Econometrics 13*(1).

Han, H. and D. Kristensen (2014). Asymptotic theory for the qmle in garch-x models with stationary and nonstationary covariates. *Journal of Business & Economic Statistics 32*(3).

Hansen, P. R. and Z. Huang (2015). Exponential garch modeling with realized measures of volatility. *Forthcoming at Journal of Business & Economic Statistics*.

Hansen, P. R., Z. Huang, and H. H. Shek (2012). Realized garch: a joint model for returns and realized measures of volatility. *Journal of Applied Econometrics 27*(6), 877–906.

Hansen, P. R., Z. Huang, and T. Wang (2015). Realized egarch, cboe vix and variance risk premium. *Working Paper*.

Hansen, P. R., A. Lunde, and J. M. Nason (2011). The model confidence set. *Econometrica 79*(2), 453–497.

Hansen, P. R., A. Lunde, and V. Voev (2014). Realized beta garch: A multivariate garch model with realized measures of volatility and covolatility. *Journal of Applied Econometrics 29.5*, 774–799.

Heber, G., A. Lunde, N. Shephard, and K. Sheppard (2009). Oxford-man institute's realized library, version 0.1.

Jensen, S. and A. Rahbek (2004). Asymptotic normality of the qmle estimator of arch in the nonstationary case. *Econometrica 72*, 641–646.

Louzis, D. P., S. Xanthopoulos-Sisinis, and A. P. Refenes (2014). Realized volatilitymodels and alternative value-at-risk prediction strategies. *Economic Modelling 40*, 101–116.

Lunde, A. and K. V. Olesen (2013). Modeling and forecasting the volatility of energy forward returns - evidence from the nordic power exchange. *Working Paper*.

Müller, U. A., M. M. Dacorogna, R. D. Davé, R. B. Olsen, O. V. Pictet, and J. E. von Weizsäcker (1997). Volatilities of different time resolutions–analyzing the dynamics of market components. *Journal of Empirical Finance 4*(2), 213–239.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics 160*(1), 246–256.

Robinson, P. M. (1995). Gaussian semiparametric estimation of long range dependence. *Annals of Statistics 23*, 1630–1661.

Shephard, N. and K. Sheppard (2010). Realising the future: forecasting with high-frequency-based volatility (heavy) models. *Journal of Applied Econometrics 25*(2), 197–231.

Shimotsu, K. (2010). Exact local whittle estimation of fractional integration with unknown mean and time trend. *Econometric Theory 26*, 501–540.

Todorova, N. (2015). The course of realized volatility in the lme non-ferrous metal market. *Economic Modelling 51*, 1–12.

Watanabe, T. S. (2012). Quantile forecasts of financial returns using realized garch models. *Japanese Economic Review 63*(1), 68–80.

Yalama, A. and S. Çelik (2013). Real or spurious long memory characteristics of volatility: Empirical evidence from an emerging market. *Economic Modelling 30*, 67–72.

Zumbach, G. (2004). Volatility processes and volatility forecast with long memory. *Quantitative Finance 4*(1), 70–86.