International Journal of Forecasting 33 (2017) 337–344

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

A vector heterogeneous autoregressive index model for realized volatility measures

Gianluca Cubadda [a,*], Barbara Guardabascio [b], Alain Hecq [c]

[a] Universita' di Roma "Tor Vergata", Dipartimento di Economia e Finanza, Via Columbia 2, 00133 Roma, Italy

[b] ISTAT, DCSC-SER/C, Viale Liegi 13, 00198 Roma, Italy

[c] Maastricht University, Department of Quantitative Economics, P.O.Box 616, 6200 MD Maastricht, The Netherlands

ARTICLE INFO

Keywords:
Common volatility
HAR models
Index models
Combinations of realized volatilities
Forecasting

ABSTRACT

This paper introduces a new model for detecting the presence of commonalities in a set of realized volatility measures. In particular, we propose a multivariate generalization of the heterogeneous autoregressive model (HAR) that is endowed with a common index structure. The vector heterogeneous autoregressive index model has the property of generating a common index that preserves the same temporal cascade structure as in the HAR model, a feature that is not shared by other aggregation methods (e.g., principal components). The parameters of this model can be estimated easily by a proper switching algorithm that increases the Gaussian likelihood at each step. We illustrate our approach using an empirical analysis that aims to combine several realized volatility measures of the same equity index for three different markets.

© 2016 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.

1. Introduction

The presence of co-movements in volatility measures is usually explained by common reactions of investors, policy makers or central banks to news relating to certain macroeconomic and financial variables. Engle and Marcucci (2006) find evidence indicating the presence of common ARCH factors (Engle & Susmel, 1993) between 435 pairs obtained from 30 stocks of the Dow Jones industrial index. However, their statistical approach might suffer from severe size distortions when applied in a multivariate setting (see Cubadda & Hecq, 2011; Hecq, Laurent, & Palm, 2016). Anderson and Vahid (2007) propose the examination of information criteria for determining the presence and number of principal component factors out of 21 Australian weekly stock return volatilities. It turns out that this latter approach is probably more robust to the presence of jumps and fat tails than the canonical correlation framework of Engle and Marcucci (2006). However, these contributions assume the dynamics of the system to be very parsimonious, contrary to the observed time series properties of daily volatility measures. For instance, the univariate heterogeneous autoregressive model (HAR; see Corsi, 2009) captures the long range dependence observed in daily time series using a restricted autoregressive model of order 22.

This paper proposes a new model for analyzing the joint behaviors of a set of daily volatility measures. We start out with a multivariate version of the HAR, namely the vector HAR (VHAR henceforth, see Bubák, Kočenda, & Žikeš, 2011). Next, we test, and consequently restrict, the VHAR by means of a multivariate autoregressive index model (Reinsel, 1983). In particular, we impose proper reduced rank restrictions on the coefficient matrices of the VHAR to obtain the vector heterogeneous autoregressive index model (VHARI henceforth).

* Corresponding author.
E-mail addresses: gianluca.cubadda@uniroma2.it (G. Cubadda), guardabascio@istat.it (B. Guardabascio), a.hecq@maastrichtuniversity.nl (A. Hecq).

http://dx.doi.org/10.1016/j.ijforecast.2016.09.002
0169-2070/© 2016 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.

The VHARI is nested within the unrestricted VHAR, which in turn is a restricted version of a vector autoregressive model (VAR) of order 22. The VHARI provides a parsimonious model whose forecasting performance can be compared with those of either less restricted multivariate models (e.g., VHAR or VAR(22)) or univariate HAR equations. At the representation theory level, the common factors obtained from the VHARI, namely the indexes, preserve the same temporal cascade structure as in the HAR; i.e., the weekly (monthly) index is equal to the weekly (monthly) moving average of the daily index. This is an important property of the VHARI that is not shared by most of the alternative aggregation methods (e.g., principal components, canonical correlations, etc.). Moreover, in a VHARI with one common component, a specification that is not rejected by the data in the empirical section of this paper, the unique index is generated by an univariate HAR model. This is not generally the case for alternative aggregation strategies either.

The rest of the paper proceeds as follows. Section 2 presents the VHAR and VHARI models, with their implications. Statistical inference is discussed in detail in Section 3. Note that we use a switching algorithm to maximize the Gaussian likelihood of a given VHARI specification. Hence, in principle, the adequacy of our set of restrictions can be checked using either information criteria or likelihood ratio tests. However, this strategy cannot be implemented for factors obtained through principal component analysis, for instance. Moreover, in the same vein as Takeuchi (1976), we propose some modified versions of the usual information criteria that are better suited for non-Gaussian series. Section 4 contains a Monte Carlo simulation exercise that documents the small-sample properties of our modelling strategy. Section 5 uses the suggested framework to combine ten realized volatility measures of the same equity index for three different markets using data from the Oxford-Man Institute of Quantitative Finance. Finally, Section 6 concludes.

2. Model representation

2.1. The vector heterogeneous autoregressive model

Our starting point for capturing the dynamic interactions within a set of n daily realized volatility measures $Y_t^{(d)} \equiv \left(Y_{1,t}^{(d)}, \ldots, Y_{n,t}^{(d)}\right)'$ is a multivariate version of the univariate HAR model (Corsi, 2009), as was used by Bubák et al. (2011) and Souček and Todorova (2013), *inter alia*.

The vector $Y_t^{(d)}$ can include either the same kind of volatility measure (e.g., the realized variance)[1] for different markets in a study of volatility co-movements or several volatility measures (realized variance, bipower variation, etc.)[2] for the same market in order to construct an optimal linear combination like that of Patton and Sheppard (2009). The latter analysis is pursued in Section 5 of this paper.

The vector heterogeneous autoregressive model (VHAR) can be written as follows:

$$Y_t^{(d)} = \beta_0 + \Phi^{(d)} Y_{t-1d}^{(d)} + \Phi^{(w)} Y_{t-1d}^{(w)} + \Phi^{(m)} Y_{t-1d}^{(m)} + \varepsilon_t, \quad t = 1, 2, \ldots, T, \tag{1}$$

where (d), (w), and (m) denote time horizons of one day, one week (five days in a week), and one month (assuming 22 days in a month) respectively, such that

$$Y_t^{(w)} = \frac{1}{5}\sum_{j=0}^{4} Y_{t-jd}^{(d)}, \qquad Y_t^{(m)} = \frac{1}{22}\sum_{j=0}^{21} Y_{t-jd}^{(d)}.$$

Here, the innovations ε_t are i.i.d. with $\mathrm{E}(\varepsilon_t) = 0$, $\mathrm{E}(\varepsilon_t \varepsilon_t') = \Sigma$ (positive definite), and finite fourth moments.

Beyond the fact that the HAR is a popular forecasting tool, two considerations that arise from our empirical analysis have led us to refer to Eq. (1) as a starting point. First, having estimated unrestricted VAR(p) models on a set of different volatility measures for each of the markets at hand, it emerges that we reject the null of no error autocorrelation for lags p of five or higher (using heteroskedasticity-robust LR tests). This means that the data have a greater dependence on the past. In principle, one could increase the VAR order considerably, but the curse of dimensionality remains a problem even when the sample size is as large as is the case in typical financial applications. Hence, Eq. (1) is a good compromise in terms of parameter proliferation, since a VAR(22) has $N^2 \times 22$ mean parameters, whereas the model in Eq. (1) needs $N^2 \times 3$ of them. Second, for the set of realized volatilities considered, the coefficient matrices $\Phi^{(d)}$, $\Phi^{(w)}$ and $\Phi^{(m)}$ are far from being diagonals, and consequently a set of individual HAR models does not seem appropriate.

The next subsection introduces additional meaningful restrictions to Eq. (1), namely the steps required to go from the VHAR to the VHARI.

2.2. The VHAR-index model

Let us further assume that Eq. (1) can be rewritten as:

$$Y_t^{(d)} = \beta_0 + \beta^{(d)} \omega' Y_{t-1d}^{(d)} + \beta^{(w)} \omega' Y_{t-1d}^{(w)} + \beta^{(m)} \omega' Y_{t-1d}^{(m)} + \varepsilon_t, \tag{2}$$

where ω is a $n \times q$ full-rank matrix. In terms of parsimony, Eq. (2) needs $4(n \times q) - q^2$ parameters instead of $n^2 \times 3$ in Eq. (1). Following Reinsel (1983), we label Eq. (2) the VHAR-index (VHARI) model. To some extent, the VHARI modeling is related to the pure variance model of Engle and Marcucci (2006), in the sense that a reduced-rank restriction is imposed on the mean parameters of a multivariate volatility model. However, one fundamental difference between Eq. (2) and the common volatility model (see also Hecq et al., 2016) stems from the fact that the former generally has a different left null space for the loading matrices of the indexes $\beta = \left[\beta^{(d)} : \beta^{(w)} : \beta^{(m)}\right]$. Obviously, common volatility is allowed in the VHARI model if there exists a full-rank $n \times s$ (with $s < q$) matrix such that $\delta' \beta = 0$.

[1] The realized covariances may also be included in $Y_t^{(d)}$, see Fengler and Gisler (2015).

[2] The realized variances are computed using $RV_t \equiv \sum_{i=1}^{M} r_{t,i}^2$, where $r_{t,i}$ are the high frequency intra-day returns, observed for M intra-day periods each day. For instance, when the market is open between 9 a.m. and 4 p.m., $M = 79$ for 5-min returns. The bipower variation $BV_t \equiv \frac{\pi}{2}\frac{M}{M-1}\sum_{i=2}^{M} |r_{t,i}||r_{t,i-1}|$ is one of the measures of the integrated volatility that is designed to be robust to jumps. See, *i.a.*, Barndorff-Nielsen and Shephard (2004) and Bauwens, Hafner, and Laurent (2012).

Beyond the important aspect in terms of parsimony, there are two further motivations for using Eq. (2). First, the indexes $f_t^{(d)} = \omega' Y_{t-1d}^{(d)}$ obtained from Eq. (2) satisfy the property

$$f_t^{(w)} = \frac{1}{5}\sum_{j=0}^{4} f_{t-jd}^{(d)}, \qquad f_t^{(m)} = \frac{1}{22}\sum_{j=0}^{21} f_{t-jd}^{(d)}, \tag{3}$$

as for the observed univariate realized volatilities. Hence, the temporal cascade structure of the HAR model is preserved, meaning that the weekly (monthly) index is equal to the weekly (monthly) moving average of the daily index. Such would not generally be the case with either traditional reduced-rank regression models, as per Engle and Marcucci (2006), or principal component methods.

Second, premultiplying both sides of Eq. (2) by ω' yields

$$f_t^{(d)} = \omega'\beta_0 + \omega'\beta^{(d)} f_{t-1d}^{(d)} + \omega'\beta^{(w)} f_{t-1d}^{(w)} + \omega'\beta^{(m)} f_{t-1d}^{(m)} + \omega'\varepsilon_t, \tag{4}$$

which shows that the indexes themselves follow a VHAR model. When $q = 1$, the unique index is generated by an univariate HAR model. This property is not shared by alternative methods of aggregating time series (e.g., averages, principal components, canonical correlations, etc.), since the resulting linear combination would generally follow a rather complicated ARMA structure; see Cubadda, Hecq, and Palm (2009), Hecq et al. (2016), and the references therein related to the final equation representation of multivariate models.

3. Statistical inference

We estimate the parameters of the model in Eq. (2) by resorting to a switching algorithm that is applied widely in cointegration analysis (see Boswijk & Doornik, 2004, and the references therein). The strategy consists of alternating between estimating ω for given values of β and Σ, and estimating β and Σ for a given value of ω. Specifically, the procedure is as follows:

1. Conditional on an (initial) estimate of ω, estimate β and Σ by OLS in Eq. (2).
2. Pre-multiplying both sides of Eq. (2) by $\Sigma^{-1/2}$, one obtains

$$\Sigma^{-1/2}(Y_t^{(d)} - \beta_0) = \Sigma^{-1/2}\beta^{(d)}\omega' Y_{t-1d}^{(d)} + \Sigma^{-1/2}\beta^{(w)}\omega' Y_{t-1d}^{(w)} + \Sigma^{-1/2}\beta^{(m)}\omega' Y_{t-1d}^{(m)} + \Sigma^{-1/2}\varepsilon_t.$$

Applying the Vec operator to both sides of the above equation and using the property $\mathrm{Vec}(ABC) = (C' \otimes A)\mathrm{Vec}(B)$, one gets

$$\begin{aligned}\mathrm{Vec}\left[\Sigma^{-1/2}(Y_t^{(d)} - \beta_0)\right] &= \left(Y_{t-1d}^{(d)\prime} \otimes \Sigma^{-1/2}\beta^{(d)}\right)\mathrm{Vec}(\omega') \\ &+ \left(Y_{t-1d}^{(w)\prime} \otimes \Sigma^{-1/2}\beta^{(w)}\right)\mathrm{Vec}(\omega') \\ &+ \left(Y_{t-1d}^{(m)\prime} \otimes \Sigma^{-1/2}\beta^{(m)}\right)\mathrm{Vec}(\omega') \\ &+ \mathrm{Vec}\left(\Sigma^{-1/2}\varepsilon_t\right),\end{aligned} \tag{5}$$

from which we can finally estimate the ω coefficients by OLS, conditional on the estimates of the parameters β and Σ obtained previously.
3. Switch between steps 1 and 2 till numerical convergence occurs.

As was shown by Boswijk (1995), the proposed switching algorithm has the property of increasing the Gaussian likelihood at each step. With respect to the Newton–Raphson method that was proposed originally by Reinsel (1983), the suggested switching algorithm has several advantages and one disadvantage. On the one hand, the switching algorithm (i) is computationally simpler; (ii) does not require any normalization condition on the parameters ω; and (iii) can be modified easily to impose over-identifying restrictions on both β and ω. On the other hand, it converges slower than Newton-type methods. Consequently, it is important to choose the initial values for the index weights ω correctly. We suggest resorting to a canonical correlation analysis between $Y_t^{(d)}$ and $\left(Y_{t-1d}^{(d)} + Y_{t-1d}^{(w)} + Y_{t-1d}^{(m)}\right)$. The canonical coefficients of the latter variable provide the Gaussian ML estimators of the elements of ω when $\beta^{(d)} = \beta^{(w)} = \beta^{(m)}$.

As was suggested by a referee, a normalization on the parameters ω can help the switching algorithm to converge to the same parameterization of Eq. (2), though it is not actually necessary. One way to achieve identification of the VHARI is to use $\omega' = [I_q, \varpi']$, where ϖ is a $(n-q) \times q$ matrix, and partition the predictors conformably as $Y_{t-1d}^{(j)} = [Y_{1,t-1d}^{(j)\prime}, Y_{2,t-1d}^{(j)\prime}]'$ for $j = d, w, m$. An alternative version of the switching algorithm that imposes such identification would then require Eq. (5) to be replaced with

$$\begin{aligned}\mathrm{Vec}\Big[\Sigma^{-1/2}(Y_t^{(d)} - \beta_0 - \beta^{(d)} Y_{1,t-1d}^{(d)} &- \beta^{(w)} Y_{1,t-1d}^{(w)} - \beta^{(m)} Y_{1,t-1d}^{(m)})\Big] \\ &= \left(Y_{2,t-1d}^{(d)\prime} \otimes \Sigma^{-1/2}\beta^{(d)}\right)\mathrm{Vec}(\varpi') \\ &+ \left(Y_{2,t-1d}^{(w)\prime} \otimes \Sigma^{-1/2}\beta^{(w)}\right)\mathrm{Vec}(\varpi') \\ &+ \left(Y_{2,t-1d}^{(m)\prime} \otimes \Sigma^{-1/2}\beta^{(m)}\right)\mathrm{Vec}(\varpi') + \mathrm{Vec}\left(\Sigma^{-1/2}\varepsilon_t\right),\end{aligned}$$

from which one can estimate the ϖ coefficients, conditional on β and Σ, by OLS.

Note that a numerical stability problem may arise when the number of series is very large. One possible solution is to resort to a properly "regularized" estimate of the autocorrelation matrix function of the series $Y_t^{(d)}$ instead of on the natural one that is used implicitly in our procedure (see Bernardini & Cubadda, 2015, for details).

The number of indexes q can be identified using the usual information criteria proposed by Schwarz (BIC), Hannan-Quinn (HQIC) and Akaike (AIC). We also propose some variants of them that are based on the theoretical framework developed by Takeuchi (1976). In short, Takeuchi extends the AIC by relaxing the strong assumption that the set of candidate models includes the true model. This extension is relevant in our case for at least two reasons. First, HAR processes are generally seen as approximations of long-memory processes (Corsi, 2009). Second, the residuals of HAR models are typically non-Gaussian and heteroskedastic (e.g., Corsi, Audrino, & Renò, 2012; Corsi, Mittnik, Pigorsch, & Pigorsch, 2008), whereas our switching algorithm aims to maximize the Gaussian likelihood. In the Appendix A, we develop a Takeuchi-type modification of the traditional information criteria for our VHARI models. We denote the modified criteria as MAIC, MHQIC, and MBIC.

4. Monte Carlo analysis

This section presents a Monte Carlo study which aims to evaluate the finite sample performances of our method. The previous section has shown how we can estimate ω using a switching algorithm for a fixed number of indexes q of the VHARI. We now investigate the relative merits of the traditional and modified information criteria for model identification, estimation, and forecasting.

The Monte Carlo design simulates demeaned realized volatilities $\overline{Y}_t^{(d)} = Y_t^{(d)} - \mathrm{E}\left(Y_t^{(d)}\right)$ that are generated by the following model:

$$\overline{Y}_t^{(d)} = A\Delta^{(d)}A^{-1}\overline{Y}_{t-1d}^{(d)} + A\Delta^{(w)}A^{-1}\overline{Y}_{t-1d}^{(w)} + A\Delta^{(m)}A^{-1}\overline{Y}_{t-1d}^{(m)} + \varepsilon_t, \quad t = 1, 2, \ldots, T, \qquad (6)$$

where A is a full-rank $n \times n$ matrix, $\Delta^{(d)}$ is a diagonal matrix with the first q diagonal elements being drawn from a $U_n(0.36, 0.399)$ and the remaining elements being equal to zero, and $\Delta^{(w)}$ and $\Delta^{(m)}$ are two diagonal matrices with the first q diagonal elements being drawn from a $U_n(0.28, 0.30)$ and the remaining elements being equal to zero. Note that Eq. (6) implies that series $\overline{Y}_t^{(d)}$ are generated by a VHARI model with parameters

$$\beta^{(j)} = A\Delta_{\bullet q}^{(j)}, \qquad \omega' = A_{q\bullet}^{-1}, \quad \text{for } j = d, w, m,$$

where $\Delta_{\bullet q}^{(j)}$ is the matrix formed by the first q columns of $\Delta^{(j)}$, and $A_{q\bullet}^{-1}$ is the matrix formed by the first q rows of A^{-1}. We reproduce the positive co-movements of the realized volatility measures that we observe in our data by generating the matrix A using a $n \times n$ half-normal distribution.

In order to both take into account the positiveness of the realized variances and reproduce the well-known volatility in volatility phenomenon (Corsi et al., 2008), the elements of ε_t have a conditional log-normal error distribution with GARCH variances (see e.g. Barndorff-Nielsen & Shephard, 2002), on the use of the log-normal distribution for realized volatilities). In particular, $\varepsilon_{it} = z_{it}\sqrt{h_{it}}$ for $i = 1, \ldots, n$, where $z_{it} = [u_{it} - \mathrm{E}(u_{it})]/\sqrt{\mathrm{Var}(u_{it})}$, $\ln(u_{it})$ is the ith element of an i.i.d. $N(0, \mathrm{I}_n)$, and $h_{it} = 0.01 + 0.25\varepsilon_{it-1}^2 + 0.74h_{it-1}$.

We generate $1000 + 200$ observations of each series, where the first 100 points are used as the burn-in period, the central $T = 1000$ observations are used for estimation, and the final points are used to compute 100 one-step-ahead forecast errors. We consider both $n = 10$ and 20 series, with $q = 1, 2, 4$ indexes. Our decisions about q and T are guided by the features of the variables that we analyze in the empirical application.

We evaluate the merits of the six information criteria by means of three statistics over 1000 replications. First, the percentage of times the estimated number of indexes $\widehat{q}$ is equal to the true index q. Second, the average mean square forecast error relative to the unrestricted VHAR forecasts (ARMSFE). Third, the Frobenius distance between the estimated mean VHARI parameters and the true ones, relative to the Frobenius distance of the OLS estimates of the mean VHAR parameters (RFD). We assess the significance of the differences in performances between the traditional and modified information criteria using the t-test for the null hypothesis that the differences between the ARMSFEs or the RFDs over the 1000 replications are centered on zero, and McNemar's test of the null hypothesis that the probabilities of identifying the true number of indexes are the same. The results are reported in Table 1.

We observe that the MBIC is the best criterion according to all three statistics when $q = 1$, and does significantly better than the BIC for $n = 10$. On the other hand, when $q = 4$, the best criteria are BIC and MHQIC. The performances of BIC, MBIC, and MHQIC are similar when $q = 2$. The AIC, its modified version, and the HQIC never perform best. Regarding the usefulness of our modifications to the traditional information criteria, AIC and HQIC uniformly perform worse than their modified counterparts, whereas matters are more controversial for the BIC. In fact, MBIC improves on BIC significantly when $n = 10$ and $q = 1$, whereas BIC outdoes MBIC significantly when $q = 4$. Interestingly, the most accurate forecasts are very often matched with the highest percentages of correct model identifications. This suggests that, if the best performing information criteria provide conflicting results, an out-of-sample forecasting exercise provides valuable information on the choice of q. We pursue this strategy in the empirical application.

5. Empirical application

This section illustrates our approach, with the aim of detecting the existence of common components within ten realized volatility measures. Following Patton and Sheppard (2009), the main idea is to build linear combinations of different volatility indicators and evaluate their merits through an out-of-sample forecasting exercise. In particular, the target variable should be an unbiased proxy for the unobserved quadratic variation (QV), whereas the predictors are (linear combinations of the) lags of the individual indicators and their linear combinations. Following Liu, Patton, and Sheppard (2015), Patton (2011) and Patton and Sheppard (2009), and for the sake of comparison with those papers, we use the daily squared open-to-close return $\tilde{r}_t^2$ as our target variable for the latent volatility measure. The daily squared return is assumed to be free from microstructure and other biases, and so is an unbiased, albeit noisy, estimator of QV. Moreover, since as our main goal is to construct a combination of realized volatility measures, we should not use one of the measures that is

Table 1
Monte Carlo results.

q		$n=10$			$n=20$		
		$\%(\widehat{q}=q)$	RFD	ARMSFE	$\%(\widehat{q}=q)$	RFD	ARMSFE
1	BIC	99.10	19.47	89.70	**100.00**	**10.50**	**79.61**
	HQIC	90.50	21.44	89.94	95.50	11.10	79.70
	AIC	17.80	51.22	93.30	19.40	34.73	83.18
	MBIC	**100.00*****	**19.30*****	**89.68****	99.90	10.51	79.62
	MHQIC	98.8***	19.58***	89.72***	99.50***	10.57***	79.62***
	MAIC	68.20***	28.88***	90.72***	65.4***	18.82***	80.80***
2	BIC	98.80	27.83	91.31	**100.00**	**14.55**	**79.98**
	HQIC	89.70	29.92	91.55	96.40	15.02	80.05
	AIC	19.20	56.29	94.45	18.70	37.18	83.65
	MBIC	**99.10**	**27.72**	91.59	99.90	14.56	80.04
	MHQIC	98.80***	27.81***	**91.29*****	99.00***	14.73***	80.01***
	MAIC	65.10***	37.44***	92.29***	61.70***	22.99***	81.18***
4	BIC	**99.30*****	**42.61*****	94.11***	98.90***	**22.15*****	82.24***
	HQIC	94.40	43.57	94.12	97.80	22.44	82.19
	AIC	27.70	62.88	96.02	21.80	41.35	84.97
	MBIC	86.10	43.17	96.87	91.30	22.32	83.18
	MHQIC	99.10***	42.71***	**94.05**	**99.00*****	22.29***	**82.17****
	MAIC	71.59***	49.24***	94.65***	55.50***	31.81***	83.46***

Note: $\%(\widehat{q}=q)$ is the percentage of times each information criterion (IC) detects the true number of indexes in the VHARI model. RFD is the Frobenius distance between the estimated and true parameters, relative to the Frobenius distance of OLS. ARMSFE is the average mean square forecast error relative to the n HAR univariate forecasts. The best result for each pair (n, q) is in bold.
** Indicates significance at the 5% level of the t-tests of equal ARMSFEs or RFDs and McNemar's test on the differences between $\%(\widehat{q}=q)$ values of the methods identified by a given IC and its Takeuchi-type version.
*** Indicates significance at the 1% level of the t-tests of equal ARMSFEs or RFDs and McNemar's test on the differences between $\%(\widehat{q}=q)$ values of the methods identified by a given IC and its Takeuchi-type version.

Table 2
Data used.

Realized volatility measure	Acronym
Realized variance (5 min.)	RV5
Realized kernel	RK
Realized variance (5 min. using 1 min. subsamples)	RV5_1
Realized variance (10 min.)	RV10
Realized variance (10 min. using 1 min. subsamples)	RV10_1
Bipower variation (5 min.)	BV5
Bipower variation (5 min. using 1 min. subsamples)	BV5_1
Median truncated realized variance	MTRV
Realized semivariance (5 min.)	RSV5
Realized semivariance (5 min. using 1 min. subsamples)	RSV5_1

Table 3
ARMSFEs and quartiles of the $\widehat{q}$ distribution: SandP500.

Method/criterion	ARMSFE			$[Q_1\ Q_2\ Q_3]$
	$h=1$	$h=5$	$h=22$	
VHARI/BIC	94.4*	**82.9***	94.2	[3 4 7]
VHARI/HQIC	94.6*	84.4*	95.8	[6 7 8]
VHARI/AIC	94.2*	84.1*	95.4	[9 10 10]
VHARI/MBIC	**92.9***	85.6*	93.4*	[1 1 1]
VHARI/MHQIC	94.3*	85.7*	**90.5***	[1 1 2]
VHARI/MAIC	94.5*	84.3*	93.8	[5 6 7]
VHAR	94.2*	84.2*	95.4	

Note: h is the forecasting horizon. ARMSFE is the average of the mean square forecast errors relative to the HAR univariate forecasts. Q_i indicates the ith quartile of the distribution of the number of indexes. The best result for each h among the multivariate methods is denoted in bold.
* Indicates models that belong to the superior set at the 20% level.

used as a predictor as our target variable, in order to avoid the risk of putting too much weight on it.

We consider daily series of the measures that are reported in Table 2, spanning the period 01/01/2000 to 10/29/2015, for three equity indexes: SandP500 for the U.S., FTSE 100 for the U.K., and the Nikkei 225 for Japan. These series are downloaded from the webpage of the Oxford-Man Institute of Quantitative Finance (see Heber, Lunde, Shephard, & Sheppard, 2009).

Before evaluating the VHARI as an aggregation method, we first check the adequacy of the VHARI restrictions. In particular, we use a rolling window of 1000 observations to compute the h-step-ahead direct forecast for $h = 1, 5, 22$ from the individual HARs, the unrestricted VHAR and the VHARI, where the number of indexes q is chosen based on both the usual information criteria and the Takeuchi-type modified ones.

Tables 3–5 report both the average mean square forecast errors relative to the HAR forecasts (ARMSFE) and the quartiles of the distribution of the numbers of indexes that are obtained using the various information criteria. An ARMSFE with a value of 100 indicates that a given multivariate model performs as well as the individual HAR models, on average over the 10 measures. We find the set of multivariate models which forecast equally well using the model confidence set (MCS) analysis of Hansen, Lunde, and Nason (2011). In particular, the test of the null hypothesis of equal predictive ability at the 20% level is implemented using a block bootstrap scheme with 5000 resamples.

The results indicate that the VHARI outperforms the univariate HARs for both the SandP500 and the Nikkei when $h = 5, 22$, whereas the reverse applies in the cases of FTSE and the Nikkei when $h = 1$. Interestingly, the unrestricted VHAR is never the best performer. The VHARI specifications suggested by the modified criteria always belong to the set of superior multivariate forecasting models, whereas the models identified by the conventional criteria fail to do so on several occasions. In the case of

Table 4
ARMSFEs and quartiles of the $\widehat{q}$ distribution: FTSE.

Method/criterion	ARMSFE			[Q1 Q2 Q3]
	$h=1$	$h=5$	$h=22$	
VHARI/BIC	108.6*	111.5*	106.6*	[3 4 4]
VHARI/HQIC	110.1	114.2	**105.1***	[5 6 6]
VHARI/AIC	108.9*	113.8*	108.2*	[8 9 10]
VHARI/MBIC	**105.8***	**106.7***	105.8*	[1 1 2]
VHARI/MHQIC	109.1*	108.4*	112.6*	[1 3 4]
VHARI/MAIC	108.5*	113.3*	109.4*	[4 4 5]
VHAR	108.8*	114.0*	107.8*	

Notes: see the notes to Table 3.

Table 5
ARMSFEs and quartiles of the $\widehat{q}$ distribution: NIKKEI.

Method/criterion	ARMSFE			[Q1 Q2 Q3]
	$h=1$	$h=5$	$h=22$	
VHARI/BIC	112.7	98.9	92.7	[1 4 5]
VHARI/HQIC	113.7	98.3	94.3	[3 5 9]
VHARI/AIC	113.9	97.6	93.3	[6 8 10]
VHARI/MBIC	**110.0***	**93.8***	**86.8***	[1 1 1]
VHARI/MHQIC	113.8	98.4	90.8	[1 4 4]
VHARI/MAIC	113.5	98.1	93.0	[4 5 6]
VHAR	113.8	97.3	92.2	

Note: see the notes to Table 3.

the Nikkei, the VHARI with q chosen by the MBIC is the unique element in the superior set. Moreover, a modified criterion, most often MBIC, identifies the VHARI model that ranks first in the superior set in seven cases out of nine. These findings are in line with the theoretical properties of the Takeuchi corrections, which take into account the heteroskedasticity and excess kurtosis that are typically present in the residuals of HAR-type models. Looking at the quartiles of the empirical distributions of the estimated q, we see that the best forecasting model is associated almost uniformly with a small number of indexes, mostly $q=1$. Overall, these empirical findings suggest that using the VHARI to build a single linear combination of the 10 volatility indicators is appropriate for the analysis that follows.

Next, we resort again to a rolling window of 1000 observations for computing the direct h-step-ahead forecasts from the models

$$\tilde{r}^2_{i,t+h} = \alpha_i + \alpha_i^{(d)} y_{i,t}^{(d)} + \alpha_i^{(w)} y_{i,t}^{(w)} + \alpha_i^{(m)} y_{i,t}^{(m)} + \epsilon_{i,t},$$

where $\tilde{r}^2_{i,t}$ is the daily squared return for each market $i=1,2,3$ and $y_{i,t}^{(j)}$ indicates each of the ten realized volatility measures for $j = d, w, m$, and those obtained from the models

$$\tilde{r}^2_{i,t+h} = \alpha_i + \alpha_i^{(d)} f_{i,t}^{(d)} + \alpha_i^{(w)} f_{i,t}^{(w)} + \alpha_i^{(m)} f_{i,t}^{(m)} + \epsilon_{i,t}, \quad (7)$$

where the predictors are scalar indexes with weights ω_i that are constructed using three alternative methods: (i) each element of ω_i is equal to 1/10 (i.e., the index at each frequency is the simple mean of the individual volatility measures at that frequency); (ii) ω_i is the eigenvector the corresponds to the first principal component of the ten daily indicators; and (iii) ω_i is estimated through the multivariate VHARI system with $q=1$. Hence, we compute thirteen forecasts of the daily squared returns for each market i and forecasting horizon h.

Table 6
RMSFE: SandP500.

Index	RMSFE		
	$h=1$	$h=5$	$h=22$
VHARI	**87.3***	100.5	99.5
PC	99.7	100.0	100.0
RV5	98.0	99.1	99.5
RK	99.5	99.9	99.9
RV5_1	100.3	100.3	100.3
RV10	95.6	**98.3**	99.6
RV10_1	100.9	100.5	99.9
BV5	104.9	101.5	100.5
BV5_1	102.1	101.0**	100.5
MTRV	102.0	102.1**	100.7
RSV5	94.1	99.3	**99.4**
RSV5_1	96.7**	99.0	100.0

Note: RMSFE is the mean square forecast error relative to the mean factor forecast. VHARI is the index produced by the proposed model with $q=1$. PC is the first principal component of the ten measures of realized volatility. See Table 2 for the remaining acronyms.
* Indicates significance at the 10% level for the Diebold–Mariano test of equal RMSFEs of a model and the benchmark. The best result is given in bold.
** Indicates significance at the 5% level for the Diebold–Mariano test of equal RMSFEs of a model and the benchmark. The best result is given in bold.

Table 7
RMSFE: FTSE.

Index	RMSFE		
	$h=1$	$h=5$	$h=22$
VHARI	**97.8**	100.0	**98.8**
PC	100.1*	100.0	100.0
RV5	100.3*	99.9	99.9
RK	99.6	99.8	99.9
RV5_1	99.7*	100.1	100.0
RV10	100.1	100.6	100.6
RV10_1	99.7	100.2	100.3
BV5	101.0	99.9	99.8
BV5_1	99.9	99.8	100.2**
MTRV	98.8	**98.4**	99.2
RSV5	101.6**	99.9	99.9
RSV5_1	100.8	100.4	100.2

Notes: see the notes to Table 6.

Tables 6–8 report the mean square forecast errors relative to the forecasts obtained using the model in Eq. (7) with uniform index weights (RMSFE henceforth). We use the simple mean index model as the benchmark because Patton and Sheppard (2009) show that it is difficult to beat in forecasting comparisons. We also report the results of Harvey, Leybourne, and Newbold's (1997) version of the Diebold and Mariano (1995) test. In particular, the null hypothesis that the MSFE of a given model is the same as that of the benchmark is tested against the alternative that the worse of the two models has a larger MSFE.

Overall, the results are as follows. For the SandP500, the VHARI indicator model performs much better than models that are based on both individual indicators and aggregates when $h=1$. The improvement over the benchmark is significant at the 10% level. For longer forecasting horizons, the performance of the VHARI factor is similar to that of the mean factor and the best individual indicator, i.e., the 5-minute realized semivariance. For the FTSE, the VHARI indicator performs best when $h=1, 22$, although the improvements over the benchmark are not significant.

Table 8
RMSFE: NIKKEI.

Index	RMSFE		
	$h = 1$	$h = 5$	$h = 22$
VHARI	103.6	100.8	**99.3**
PC	100.1	99.9	100.0
RV5	100.4	**99.4**	100.0
RK	**96.0****	99.8	100.1
RV5_1	97.1	99.5	99.9
RV10	102.8	101.7*	100.4
RV10_1	98.3	100.6	100.1
BV5	102.3	99.9	99.8
BV5_1	98.9	100.5	100.1
MTRV	96.3	101.8	99.6
RSV5	101.7	100.8	100.0
RSV5_1	100.2	101.2	100.1

Notes: see the notes to Table 6.

When $h = 5$, the best performer is the model based on the median truncated realized variance. For the Nikkei, the realized kernel measure is significantly superior to the benchmark when $h = 1$, whereas the models based on the 5-minute realized variance and the VHARI factor are the best performers when $h = 5$ and $h = 22$, respectively.

Finally, we can conclude from this illustration that the VHARI factor model performs best in one-third of cases in this example. It is very often superior to models based on the simple mean or the first principal component. Moreover, there are no individual indicators that performs systematically better than the competitors across different markets and for different forecast horizons. These findings suggest that it may be worthwhile to add the VHARI to the multivariate realized volatility modelling toolkit.

6. Conclusions

This paper has proposed the VHARI model, a multivariate generalization of the HAR model of Corsi (2009), which allows for the parsimonious modelling of a vector of realized volatilities. In particular, the realized volatility measures can be explained as linear functions of a few indexes, which preserve the same temporal cascade structure as the autoregressive terms of the univariate HAR model. The parameters of the VHARI model can be estimated by means of a switching algorithm that increases the Gaussian likelihood at each step. Based on the work of Takeuchi (1976), we have modified traditional information criteria by allowing for non-Gaussianity and heteroskedasticity in the errors of our model. Finally, we have illustrated the practical value of the proposed methods by means of an empirical application to a set of ten realized volatility measures for the SandP500, FTSE and the Nikkei equity indexes.

Acknowledgments

The authors thank Sébastien Laurent, a referee, an associate editor, and participants at our presentations at the Sixth Italian Congress of Econometrics and Empirical Economics, Salerno, January 2015, the 8th CSDA International Conference on Computational and Financial Econometrics, Pisa, December 2014, and the workshop "Nouveaux développements dans la modélisation et la prévision des risques extrêmes en finance", Marseille, May 2015. The usual disclaimers apply.

Appendix A

Assume that the candidate models have the form

$$Y_t = \theta' X_t + \varepsilon_t, \tag{8}$$

where Y_t and X_t are random variables of dimensions n and k respectively, θ is an $n \times k$ coefficient matrix, and ε_t are i.i.d. $N(0, \Sigma)$ errors that are distributed independently of X_t, whereas the true model is

$$Y_t = \mathrm{E}(Y_t|X_{0,t}) + \epsilon_t,$$

where $X_{0,t}$ is a random variable of dimension $k_0 \leq k$, such that its elements are a subset of those of X_t, and ϵ_t are i.i.d. non-normal errors with $\mathrm{E}(\epsilon_t) = 0, \mathrm{E}(\epsilon_t\epsilon_t') = \Sigma_0$, and finite fourth moments.

If the model in Eq. (8) is estimated by Gaussian ML, the penalty term of the AIC is not fully appropriate, because its derivation is based on the assumption that the set of candidate models includes the true model; see e.g. Burnham and Anderson (2002). Takeuchi (1976) relaxed this assumption and obtained the following criterion:

$$\mathrm{TIC} = \ln(|\widehat{\Sigma}|) + 2\widehat{\eta}/T,$$

where

$$\widehat{\eta} = \sum_{t=1}^{T} \widehat{\varepsilon}_t' \widehat{\Sigma}^{-1} \widehat{\varepsilon}_t h_{tt} + \frac{1}{2}\left[T^{-1}\sum_{t=1}^{T}\left(\widehat{\varepsilon}_t'\widehat{\Sigma}^{-1}\widehat{\varepsilon}\right)^2 - n(n+2)\right], \tag{9}$$

$h_{tt} = X_t'(\sum_{i=1}^{t} X_t X_t')^{-1} X_t$, $\widehat{\varepsilon}_t$ are the OLS residuals, and $\widehat{\Sigma}$ is the residual covariance matrix; see Yanagihara (2006) for further details.

Note that if ϵ_t are i.i.d. Gaussian errors and the sample size T is large, the first term on the right-hand side of Eq. (9) will be centered on nk,[3] whereas the second term will be centered on zero.[4] Hence, $\widehat{\eta} \simeq nk$, namely the number of free parameters in θ, as in the AIC.

However, when using Takeuchi's framework to identify the number of indexes q in the VHARI model in Eq. (2), it is necessary to take into account the fact that the number of free mean parameters is equal to $4nq - q^2$ instead of $3nq$, which would be the large sample mean of $\widehat{\eta}$ under Gaussianity. Hence, we propose the use of

$$\widetilde{\eta} = \widehat{\eta} + q(n - q)$$

in place of $4nq - q^2$ in the formulae of the traditional information criteria. This leads to the definition of the following modified information criteria:

$$\mathrm{MAIC} = \ln(|\widehat{\Sigma}|) + 2\widetilde{\eta}/T,$$
$$\mathrm{MHQIC} = \ln(|\widehat{\Sigma}|) + 2\widetilde{\eta}\ln(\ln(T))/T,$$
$$\mathrm{MBIC} = \ln(|\widehat{\Sigma}|) + \widetilde{\eta}\ln(T)/T,$$

which are robust to the presence of heteroskedasticity and excess kurtosis.

3 This result follows from the use of the law of iterated expectations, noting that $\sum_{t=1}^{T} h_{tt}$ is the trace of the projection matrix of the variables' Xs.

4 This result follows from the fact that the term in square brackets in Eq. (9) is a consistent estimator of multivariate kurtosis.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.ijforecast.2016.09.002.

References

Anderson, H., & Vahid, F. (2007). Forecasting the volatility of Australian stock returns: do common factors help? *Journal of Business and Economic Statistics*, *25*, 76–90.

Barndorff-Nielsen, O. E., & Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B*, *64*, 253–280.

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

Bernardini, E., & Cubadda, G. (2015). Macroeconomic forecasting and structural analysis through regularized reduced-rank regression. *International Journal of Forecasting*, *31*, 682–691.

Bauwens, L., Hafner, Ch., & Laurent, S. (2012). Volatility models. In L. Bauwens, Ch. Hafner, & S. Laurent (Eds.), *Handbook of volatility models and their applications*. Wiley, Ch. 1.

Boswijk, H.P. (1995). Identifiability of cointegrated systems. Tinbergen Institute Working Paper, 95/78.

Boswijk, H. P, & Doornik, J. A. (2004). Identifying, estimating and testing restricted cointegrated systems: An overview. *Statistica Neerlandica*, *58*, 440–465.

Bubák, V., Kočenda, E., & Žikeš, F. (2011). Volatility transmission in emerging European exchange markets. *Journal of Banking and Finance*, *35*, 2829–2841.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: a practical information-theoretic approach*. New York: Springer-Verlag.

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

Corsi, F., Audrino, F., & Renò, R. (2012). HAR modeling for realized volatility forecasting. In L. Bauwens, Ch. Hafner, & S. Laurent (Eds.), *Handbook of volatility models and their applications*. Wiley, Ch. 15.

Corsi, F., Mittnik, S., Pigorsch, C., & Pigorsch, U. (2008). The volatility of realized volatility. *Econometric Reviews*, *27*, 46–78.

Cubadda, G., & Hecq, A. (2011). Testing for common autocorrelation in data rich environments. *Journal of Forecasting*, *30*, 325–335.

Cubadda, G., Hecq, A., & Palm, F. C. (2009). Studying co-movements in large multivariate models prior to modeling. *Journal of Econometrics*, *148*, 25–35.

Diebold, F. X., & Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *13*, 253–263.

Engle, R. F., & Marcucci, J. (2006). A long-run pure variance common features model for the common volatilities of the Dow Jones. *Journal of Econometrics*, *132*, 7–42.

Engle, R. F., & Susmel, R. (1993). Common volatility in international equity markets. *Journal of Business and Economic Statistics*, *11*, 167–176.

Fengler, M., & Gisler, K. (2015). A variance spillover analysis without covariances: What do we miss? *Journal of International Money and Finance*, *51*, 174–195.

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

Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, *13*, 281–291.

Heber, G., Lunde, A., Shephard, N., & Sheppard, K. (2009). *Oxford-Mann institute's realized library version 0.1*. University of Oxford: Oxford-Man Institute.

Hecq, A., Laurent, S., & Palm, F. (2016). On the univariate representation of BEKK models with common factors. *Journal of Time Series Econometrics*, *8*, 91–113.

Liu, L., Patton, A., & Sheppard, K. (2015). Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. *Journal of Econometrics*, *187*, 293–311.

Patton, A. (2011). Data-based ranking of realised volatility estimators. *Journal of Econometrics*, *161*, 284–303.

Patton, A., & Sheppard, K. (2009). Optimal combinations of realised volatility estimators. *International Journal of Forecasting*, *25*, 218–238.

Reinsel, G. (1983). Some results on multivariate autoregressive index models. *Biometrika*, *70*, 145–156.

Souček, M., & Todorova, N. (2013). Realized volatilities transmission between crude oil and equity futures markets: A multivariate HAR approach. *Energy Economics*, *40*, 586–597.

Takeuchi, K. (1976). Distribution of informational statistics and a criterion of model fitting. *Suri-Kagaku (Mathematical Sciences)*, *153*, 12–18. in Japanese.

Yanagihara, H. (2006). Corrected version of AIC for selecting multivariate normal linear regression models in a general nonnormal case. *Journal of Multivariate Analysis*, *97*, 1070–1089.

Gianluca Cubadda is Professor of Economic Statistics at the University of Rome "Tor Vergata", Italy. His main research interests are time series analysis and econometrics. His articles have appeared in various journals, including *Computational Statistics & Data Analysis*, *Econometric Reviews*, *Economic Modelling*, *Economics Letters*, *Empirical Economics*, *International Journal of Forecasting*, *Journal of Applied Econometrics*, *Journal of Econometrics*, *Journal of Forecasting*, *Journal of Time Series Analysis*, *Macroeconomic Dynamics*, and *Oxford Bulletin of Economics and Statistics*.

Barbara Guardabascio is a researcher at the National Statistical Institute (ISTAT) in Rome. Her main research interests are macroeconomic forecasting and time series analysis. Her articles have appeared in various journals, including *Economic Modelling* and *Journal of Forecasting*.

Alain Hecq is Professor of Applied Econometrics at Maastricht University, School of Business and Economics, Department of Quantitative Economics. His main research areas are time series econometrics and applied econometrics. His articles have appeared in various journals, including *Journal of Econometrics*, *Econometric Reviews*, *International Journal of Forecasting*, *Oxford Bulletin of Economics and Statistics*, *Journal of Time Series Econometrics*, *Journal of Financial Econometrics*, *Journal of Forecasting*, and *Economics Letters*.