Forecasting Realized Volatility Measures with Multivariate and Univariate Models: The Case of the US Banking Sector

Gianluca Cubadda, Alain Hecq and Antonio Riccardo
This paper compares the forecasting performances of both univariate and multivariate models for realized volatilities series. We consider realized volatility measures of the returns of 13 major banks traded in the NYSE. Since our variables are characterized by the presence of long range dependence, we use several modelling approaches that are able to capture such feature. We look at the forecasting accuracy of the considered models to make inference on the underlying mechanism that has generated volatilities of the assets. Our main conclusion is that the contagion effect among the considered volatilities is small or, at least, not well captured by the considered multivariate models.

JEL: C32

Keywords: Common volatility, long-memory processes, HAR models, index models, forecasting.
1 Introduction

Realized volatility measures, such as the realized variance (RV) or the bipower variation, are estimates of asset volatilities within a short period, for instance one day, using intra-day returns. The 5-minute Realized Variance (RV5), a benchmark often considered in empirical finance (see Liu et al., 2015), is obtained as $RV_{5t} \equiv \sum_{j=1}^{M} r_{t,j}^2$, where $r_{t,j}$ are the high frequency returns, observed for $M$ intra-day 5-min periods. Both computational simplicity and theoretical foundations make realized volatility measures very attractive among practitioners and academics for modelling time varying volatilities and monitoring financial risk. When co-volatilities are involved as well, realized volatility measures are alternatives to multivariate GARCH models for building portfolios. Andersen and Benzoni (2009) or McAleer and Medeiros (2008) are frequently cited papers introducing this subject. Since high frequency returns are difficult to obtain and tedious to manipulate and clean, several institutions, including the Oxford-Man Institute of Quantitative Finance, preprocess the data and provide various volatility measures for aggregate stock price indexes. In this paper, we instead analyze and forecast the realized volatilities of the asset prices of thirteen major US banks. We build both RV5 and median RV5 using raw intra-day data from the Thomson Reuters database. We focus on these assets since it is commonly believed that the banking sector is highly exposed to contagion, in particular since the financial crisis (Bilio et al., 2012; Ahelegbey et al., 2016). Working instead with country or sector indices might lead to confusing results as those aggregate series merge different behaviors in individual asset volatilities.

A common empirical characteristic of realized volatility measures is the presence of long-memory dependence, as suggested by the typical slow decay pattern of the empirical autocorrelation function. This feature has been documented by many authors, including Andersen et al. (2001) for exchange rates, and Corsi et al. (2008) as well as Hillebrand and Meideiros (2016) for stock prices. Although long-memory processes are also observed in other fields than financial econometrics (see Baillie, 1996 and references therein) and that several models of long range dependence have been proposed in the literature (Haldrup and Vera-Valdés, 2017), the fractional integration process of order $d$, denoted $I(d)$, has been extensively studied in econometrics and statistics since, at least, Granger (1980) and Granger and Joyeux (1980). An example of an $I(d)$ process is the fractional white noise (FWN)

\[(1 - L)^d y_t = \varepsilon_t \iff y_t = (1 - L)^{-d} \varepsilon_t,\]

with

\[(1 - L)^d = 1 - dL + \frac{d(d - 1)}{2!}L^2 - ...\]

where $L$ denotes the lag operator, $-0.5 < d < 0.5$ and $\varepsilon_t$ is a white noise sequence. For $0 < d < 0.5$, the process is long-memory with positive autocorrelation decaying at a hyperbolic rate. For $-0.5 < d < 0$, the sum of absolute values of the autocorrelations tends to a constant and the process is said to be antipersistent. The class of fractionally integrated processes extends to the autoregressive fractionally integrated moving average (ARFIMA) processes of orders $(p, d, q)$, where $\varepsilon_t$ admits a covariance stationary and invertible ARMA representation. Corsi (2009) proposed the
univariate Heterogeneous Autoregressive model (HAR) as an alternative way to approximate the long range dependence observed in volatility series. For daily series, HAR is a parsimonious restricted autoregressive model of lag order 22 with daily, weekly and monthly effects. The HAR model can easily be estimated by ordinary least squares (OLS) and it has been illustrated to perform well in forecasting exercises (see e.g. Santos and Ziegelmann, 2014, and the references therein).

The literature on the sources of long-memory is quite large, among which the aggregation across heterogeneous series argument raised by Granger (1980), the impact of structural changes that spuriously lead to the detection of a fractional integrated process (Diebold and Inoue, 2001), the linear approximation of an underlying nonlinear process (Miller and Park, 2010) or the learning by economic agents in forward looking models of expectations (Chevillon and Mavroeidis, 2017).

In this paper we focus on the alternative explanation developed in Chevillon, Hecq and Laurent (2018, CHL18). Indeed, CHL18 investigate the mechanisms underlying the long-memory feature generated from a vector autoregressive model (VAR). They start by assuming that the dynamic interactions between \( n \) daily realized volatility measures is generated by a VAR(1). Then for \( n \to \infty \), namely when the number of series increases without bounds, and under some regularity conditions that are commonly satisfied by real data, CHL18 prove that the marginal model of each individual time series (i.e. the final equation representation) is a FWN processes with the same \( d \) parameter. Having the same \( d \) parameter for different assets is in accordance for instance with Andersen et al. (2001) who found \( d = 0.4 \) for most exchange rate realized volatilities. CHL18 provide two specific examples of those conditions among which the VAR(1) coefficient matrix has diagonal elements converging to \( 1/2 \) as \( n \to \infty \), and vanishing off-diagonal elements. This means in practice that there exist contagion effects, that individually are tiny but jointly potentially important.

However, many papers have also documented the presence of co-movements in the volatility of asset returns. In integrated markets, common factors in volatility are the result of a common reaction of investors, policy makers or central banks to news/shocks in some macroeconomic and financial variables (see, inter alia, Engle et al., 1990, Diebold and Nerlove, 1989). Nevertheless, an important assumption underlying the results by CHL18 is the existence of non-zero diagonal elements in the VAR coefficient matrix along with small but not null off-diagonal ones. This would contradict the presence of a reduced-rank structure in the VAR, and hence of a particular form of commonalities, named common features in volatility, observed, inter alia, by Engle and Marcuscu (2006), Engle and Susmel (1993), Hecq et al. (2016), and Anderson and Vahid (2007). Discriminating between these two views on the cause of the contagion in a potentially high dimensional setting is challenging and it might well be unfeasible using conventional testing approaches. Indeed, we would face an obvious curse of dimensionality issue when modelling a large system. One may argue that shrinkage techniques leading to sparse regression models could be an attractive option for the problem at hand. However, the two alternative views that we want to evaluate generally imply the use of models that are non-nested each other. Moreover, the thirteen series that we consider might be seen as a subset of a very large dimensional multivariate process that we do not observe nor estimate. For instance
a very large VAR with all the NYSE asset volatilities might in principle have generated the thirteen series considered here. Hence, estimating a marginal VAR model for the considered series only and testing for the presence of contagion effects would not provide a convincing answer to the question that we wish to address in this paper, namely the origin of the contagion in volatility.

Consequently, this paper compares the forecasting performances of two different modeling strategies: on the one hand, we consider a set of univariate models potentially derived from a huge system with hidden correlations as proposed in CHL18; on the other hand, we rely on medium scale multivariate models, possibly with common factors. For the former framework, we model the long-memory feature of the individual series using both the maximum likelihood (ML) estimation of the FWN process, i.e., the ARFIMA(0,d,0) model, as well as the OLS estimates of HAR models. For the second multivariate strategy, we must be able to capture the long-memory features observed in the series as a VAR(1), with or without a reduced rank structure, does not have such a feature.

Consequently, we first look at a multivariate version of the HAR model called the Vector HAR (VHAR), introduced by Bubak et al. (2011), in order to incorporate the possibility to have long memory features. Then, we look at the presence of co-movements in volatilities and we study the performance of the VHAR Index model (VHARI) proposed by Cubadda et al. (2017), in which the VHAR is restricted by imposing a common index structure. Note that we use these multivariate models and not, for instance, generic factor models based on principal component analysis. There are two reasons for that. First the VHARI is nested within the unrestricted VHAR, which is in turn restricted versions of a VAR with 22 daily lags. Hence, the restrictions underlying the VHAR and the VHARI could, in principle, be tested for, whereas the factor structure is typically postulated in dynamic factor models. Second, at the representation theory level, the common factors obtained from the VHARI preserve the same temporal cascade structure as in the univariate HAR with the weekly (monthly) index being equal to the weekly (monthly) moving average of the daily index. This is an important property that is not shared by most of the alternative factor methods (e.g., principal components, canonical correlations, etc.) and makes the forecasts from a set of univariate models and a multivariate system easily comparable. To some extent, our paper is close to the recent work of Bauwens et al. (2018). However, the authors of this paper examine a set of 50 realized volatilities by means of both univariate models and high dimensional VARs under the constraints that are derived from CHL18. They consequently compare the forecasting performances of the different models in a nested perspective with the aim to empirically corroborate the theoretical results of CHL18. Instead, we do not impose any theoretically based constraint to the time series models that we use in our forecasting exercise.

The rest of the paper is as follows. Section 2 sketches the results on the final equation representation as well as the main conclusions of CHL18 for the explanation of the origins of the long-memory. Section 3 reviews and discusses the VHAR and the VHARI models. Section 4 motivates our study by first looking at the presence of long-memory features in the volatility in the return series of the thirteen major US banks. We further examine the properties of our series by estimating univariate
HAR and ARFIMA models as well as a VAR on the whole sample size. Section 5 compares the forecasting performances of the different models using a rolling sample scheme. Section 6 draws conclusions.

2 The univariate implications of a multivariate structure

In this Section we briefly review the main results on the marginal model of each element of a VAR process, and those of CHL18 regarding the relation between the dimension of the VAR and the presence of long-memory in the individual time series.

2.1 The final equation representation

CHL18 investigate the mechanisms underlying the long-memory feature generated from a VAR model based on the final equation representation of a multivariate \( n \)-dimensional time series \( Y_t \equiv (Y_{1t}, \ldots, Y_{nt})' \). In order to brush up some of the basic tools needed for the understanding of the paper, let us consider a VAR(1) model

\[
(I_n - A_n L) Y_t = \epsilon_t, \tag{3}
\]

where the innovations \( \epsilon_t \) are i.i.d. with \( \text{E}(\epsilon_t) = 0 \), and \( \text{E}(\epsilon_t \epsilon'_t) = \Sigma \) (positive definite).

The final equation representation (FER) is obtained by premultiplying both sides of (3) by the adjoint matrix of the matrix polynomial \( (I_n - A_n L) \):

\[
\text{adj} (I_n - A_n L) = \det (I_n - A_n L) (I_n - A_n L)^{-1},
\]

which leads to

\[
\det (I_n - A_n L) Y_t = \text{adj} (I_n - A_n L) \epsilon_t. \tag{4}
\]

It is observed in (4) that each elements of \( Y_t \equiv (Y_{1t}, \ldots, Y_{nt})' \) follows an ARMA\((n, n-1)\) process with a common autoregressive polynomial for each series. More generally, one would obtain ARMA\((np, (n-1)p)\) processes for a VAR\((p)\). Additional details about the FER can be found in Zellner and Palm (1974, 1975, 2004), Palm (1977), Cubadda et al. (2009) and Hecq et al. (2012) among others.

As a numerical example, consider the following trivariate VAR(1) model

\[
\begin{bmatrix}
  Y_{1t} \\
  Y_{2t} \\
  Y_{3t}
\end{bmatrix} =
\begin{bmatrix}
  a_{11} & a_{12} & a_{13} \\
  a_{21} & a_{22} & a_{23} \\
  a_{31} & a_{32} & a_{33}
\end{bmatrix}
\begin{bmatrix}
  Y_{1t-1} \\
  Y_{2t-1} \\
  Y_{3t-1}
\end{bmatrix} +
\begin{bmatrix}
  \epsilon_{1t} \\
  \epsilon_{2t} \\
  \epsilon_{3t}
\end{bmatrix}. \tag{5}
\]

Computing both \( \det (I_3 - A_3 L) \) and \( \text{adj} (I_3 - A_3 L) \) leads to the observation that all univariate elements of \( Y_t \) follow and ARMA\((3, 2)\) processes.
Note that the orders derived from the FER are maximal orders. There are cases, in which orders do not depend on $n$, with smaller numbers. The first important one is when $A_n$ is exactly diagonal, i.e. when

$$A_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}$$

leading to ARMA(1,0) processes for every series.

Another interesting case is when $A_n$ is of reduced rank. For instance each series is an ARMA(1,1) whatever the dimension $n$ when $\text{rank}(A_n) = 1$ such as in the example

$$A_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$$

(see Cubadda et al. 2009, for details).

2.2 High dimensional VARs and long-memory

The approach of CHL18 (2018) builds on the previous FER. They start from the same VAR(1) as in (3) but they look at the behavior of (4) when more and more data are collected. From the previous results and without root cancellation between the determinant and the adjoint, we know that when $n \to \infty$, each series is an ARMA($\infty$, $\infty$). Obviously a smaller number of significant lags are found in practice. This process with an infinite number of lags is still of short memory however, with the usual exponential decrease of the coefficient parameters.

Then for $n \to \infty$, namely when the number of series increases without bounds, and under some regularity conditions on $A_n$ that are commonly satisfied by real data, CHL18 show that the marginal model of each individual time series is a FWN process $(1 - L)^d Y_t = \epsilon_t$ with the same $d$ parameter. Having the same $d$ parameter for different assets is in accordance with Andersen et al. (2001) who found $d = 0.4$ for most exchange rate realized volatilities. Contrarily to the exactly diagonal case illustrated in the previous subsection, CHL18 consider an $n$-variate autoregressive matrix that is near diagonal such as

$$A_n = \begin{bmatrix} a & o(1) \\ o(1) & A_{n-1} + o(1) \end{bmatrix}$$

Indeed, it is frequently observed when estimating multivariate volatility models, for instance a multivariate GARCH (see e.g. Bauwens et al., 2006, and the references therein) or a BEKK (Baba et al, 1989), that the most important contribution for the volatility comes from their own lags although individually small, the off-diagonal elements are usually jointly significantly different from zero. This is intuitively what the structure of the matrix (8) captures. Note that a sparse modelling approach such as Lasso would likely force many of those small off-diagonal elements to be exactly
equal to zero.\(^1\)

With this particular framework, the FER for the first row \((Y_{1t})\) is

\[
\det (I_n - A_n L) Y_{1t} = \det (I_{n-1} - A_{n-1} L) \epsilon_{1t} + o_p(1) \tag{9}
\]

then the moving average lag polynomial associated with \(Y_{1t}\) is asymptotically

\[
Y_{1t} \approx \frac{\det (I_{n-1} - A_{n-1} L)}{\det (I_n - A_n L)} \epsilon_{1t}. \tag{10}
\]

CHL18 further parameterize \(A_n\) by defining a scalar sequence \((\delta_n)\) with \(\delta_n \in (0, 1)\) such that

\[
\lim_{n \to \infty} \delta_n = \delta \in (0, 1), \text{ and a circulant matrix } C_n \text{ such that the polynomials } \det (I_n - A_n z) \sim \det (I_n - C_n z) \text{ as } n \to \infty. \text{ } C_n \text{ is assumed to possess close to a fraction } n\delta \text{ of unit eigenvalues. The first Szegö theorem is then used to prove that, under some high level assumptions, } \det (I_{n-1} - A_{n-1} z) / \det (I_n - A_n z) \to (1 - z)^{-\delta} \text{ as } n \to \infty. \text{ From this result, it follows the weak convergence of the process } Y_{1t} \text{ to the fractional white noise } (1 - L)^{-\delta} \epsilon_{1t}. \text{ This is also the case for each of the } n \text{ series in } Y_t. \]

Those high level assumptions needed by CHL18 are satisfied for at least two specific examples of VAR(1) models. In the first example of parameterization of interest for this paper, \(A_n\) denotes a Toeplitz matrix with diagonal elements converging to \(\delta = 1/2\) as \(n \to \infty\), and with vanishing off-diagonal elements. Importantly, the off-diagonal elements decrease at an \(O(n^{-1})\) rate and the sum of each row equals 1 at all \(n\). Then as \(n \to \infty\), each series of the system behaves as an ARFIMA(0,1/2,0). Note that this result cannot hold when \(A_n\) is exactly diagonal since \(A_{n-1}\) contains \(n\delta \) or \(n\delta - 1\) unit eigenvalues, not \(\lfloor (n - 1)\delta \rfloor\). From a practical view point (see also the simulation study in CHL18) this implies having a near diagonal matrix with some contagion effects, that individually are tiny but jointly potentially important. For instance with \(n = 500\) series, there is a large value close to 0.5 (say 0.5 - \(\varepsilon\)) on the diagonal and in each rows of \(A_n\), the sum of the 499 contagion parameters are 0.5 + \(\varepsilon\).

For each series, univariate ARFIMA models can be estimated. Several estimators of the long-memory parameters of series have been proposed in the literature, including the log periodogram regression of Geweke and Porter-Hudak (1983), the Local Whittle Likelihood Estimator of Robinson (1995), as well as the usual Gaussian ML estimator of the ARFIMA\((p, d, q)\) model.

3 The Vector Heterogeneous Autoregressive model and its index extension

Corsi (2009) proposed the HAR as an alternative way to approximate the long-memory feature. Let us start with considering the univariate HAR specification. For daily series, HAR is a parsimonious restricted autoregressive model of lag order 22 with daily, weekly and monthly effects. For a

\(^{1}\text{See, for instance, Chapter 3 in Hastie et al. (2011) for an introduction to Lasso and other regularization schemes in linear regression.}\)
daily volatility measure $Y_{it}^{(day)}$ one can run by OLS for $i = 1, \ldots, n$ volatility returns the following regressions

$$Y_{it}^{(day)} = \beta_{i0} + \varphi_{i}^{(day)}Y_{i,t-1}^{(day)} + \varphi_{i}^{(w)}Y_{i,t-1}^{(w)} + \varphi_{i}^{(m)}Y_{i,t-1}^{(m)} + \varepsilon_{it}, \quad t = 1, 2, \ldots, T; \quad (11)$$

where $(day)$, $(w)$, and $(m)$ respectively denote time horizons of one day, one week (5 days a week), and one month (assuming 22 days within a month) such that

$$Y_{it}^{(w)} = \frac{1}{5} \sum_{j=0}^{4} Y_{i,t-j}^{(day)}, \quad Y_{it}^{(m)} = \frac{1}{22} \sum_{j=0}^{21} Y_{i,t-j}^{(day)}, \quad (12)$$

and where $\beta_{i0}, \varphi_{i}^{(day)}, \varphi_{i}^{(w)}, \varphi_{i}^{(m)}$ are scalar parameters; $\varepsilon_{it}$ is i.i.d. with mean 0 and variance $\sigma_{it}^2$. Note that this can be extended to time varying GARCH type errors (see the HAR-GARCH in Corsi et al., 2012).

The Vector Heterogeneous Autoregressive model (VHAR), proposed by Bubá et al. (2011) and Soucek and Todorova (2013), is a multivariate generalization of the previous univariate process (11). The VHAR reads as follows for the levels of the $n$ volatility daily series $Y_{i}^{(day)} \equiv (Y_{1}^{(day)}, \ldots, Y_{n}^{(day)})'$:

$$Y_{t}^{(day)} = \beta_{0} + \Phi_{t}^{(day)}Y_{t-1}^{(day)} + \Phi_{t}^{(w)}Y_{t-1}^{(w)} + \Phi_{t}^{(m)}Y_{t-1}^{(m)} + \epsilon_{t}, \quad t = 1, 2, \ldots, T; \quad (13)$$

with

$$Y_{t}^{(w)} = \frac{1}{5} \sum_{j=0}^{4} Y_{t-j}^{(day)}, \quad Y_{t}^{(m)} = \frac{1}{22} \sum_{j=0}^{21} Y_{t-j}^{(day)}, \quad (14)$$

and where $\beta_{0}$ is an $n \times 1$ vector of intercepts, and $\Phi_{t}^{(day)}, \Phi_{t}^{(w)}, \Phi_{t}^{(m)}$ are $n \times n$ coefficient matrices.

Clearly the VHAR in (13) involves $(n^2 \times 3)$ parameters in addition to the vector of intercepts and is therefore much more parsimonious than a VAR with 22 unrestricted daily lags. Moreover, a VHAR is an interesting model to consider, as it is able to generate long-memory features and to introduce contagion effects between volatilities. System (13) can easily be estimated by multivariate least square regressions, which means using OLS equation by equation if no cross equation restrictions are present.

Let us further assume that (13) can be rewritten as follows

$$Y_{t}^{(day)} = \beta_{0} + \beta^{(day)} \omega^{t} Y_{t-1}^{(day)} + \beta^{(w)} \omega^{t} Y_{t-1}^{(w)} + \beta^{(m)} \omega^{t} Y_{t-1}^{(m)} + \epsilon_{t}, \quad (15)$$

where $\omega$ is a full-rank $n \times q$ matrix, and $\beta^{(day)}, \beta^{(w)}, \beta^{(m)}$ are $n \times q$ coefficient matrices. Since we can always normalize $\omega$ such as $\omega^{t} = (I_q, \varpi^t)$, where $\varpi$ is a $(n - q) \times q$ matrix, model (15) needs $4(n \times q) - q^2$ parameters instead of $n^2 \times 3$ of them as in (13). Following Reinsel (1983), we label (15) as the VHAR-index (VHARI) model.

\footnote{We use the notation $day$ to avoid the confusion with the long-memory parameter $d$.}
Beyond the important aspect in terms of parsimony that is shared with many factor models, there are two further motivations for using (15). First, the indexes $f_t^{(day)} = \omega' Y_{t-1}^{(day)}$ obtained from (15) satisfy the property

$$f_t^{(w)} = \frac{1}{5} \sum_{j=0}^{4} f_{t-j}^{(day)}, \quad f_t^{(m)} = \frac{1}{22} \sum_{j=0}^{21} f_{t-j}^{(day)},$$  

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. This would not be generally the case with either traditional reduced-rank regression models as in Engle and Marcucci (2006) or principal component methods.

Second, premultiplying both sides of (15) by $\omega'$ yields

$$f_t^{(day)} = \omega' \beta_0 + \omega' \beta^{(day)} f_{t-1}^{(day)} + \omega' \beta^{(w)} f_{t-1}^{(w)} + \omega' \beta^{(m)} f_{t-1}^{(m)} + \omega' \epsilon_t,$$

which shows that the indexes themselves follow a VHAR model. Moreover, when $q = 1$, the unique index is generated by an univariate HAR model. This property is not shared by alternative methods to aggregate 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 et al., 2009; Hecq et al., 2016, and the references related to the final equation representation of multivariate models therein).

To some extent, the VHARI model is related to the pure variance model of Engle and Marcucci (2006), in the sense that a reduced-rank restriction is imposed to the mean parameters of a multivariate volatility model. However, a fundamental difference between (15) and the common volatility model (see also Hecq et al., 2016) stems from the fact that the former has in general a different left null space for the loading matrices of the factors $\beta = (\beta^{(day)}, \beta^{(w)}, \beta^{(m)})$. Obviously, common volatility is allowed in the VHARI model as well in the case that the loading matrix $\beta$ has reduced column rank.

In order to better appreciate the differences between the VHARI and a VHAR with a reduced-rank structure, let us write the latter as

$$Y_t^{(day)} = \beta_0 + \alpha \psi^{(day)} Y_{t-1}^{(day)} + \alpha \psi^{(w)} Y_{t-1}^{(w)} + \alpha \psi^{(m)} Y_{t-1}^{(m)} + \epsilon_t, \quad t = 1, 2, ..., T,$$

where $\alpha$ is full-rank $n \times q$ matrix, and $\psi^{(day)}, \psi^{(w)}, \psi^{(m)}$ are $n \times q$ coefficient matrices.

Reduced-rank VAR models are popular in economics and finance because of their interpretation in terms of common features and their ease in estimation (see e.g. Centoni and Cubadda, 2015, and the references therein). However, although Equations (15) and (18) look similar, the interpretation of the latter is less intuitive since the coefficients of the common components, namely $(\psi^{(day)}, \psi^{(w)}, \psi^{(m)})$, are specific for each frequency. Hence, Model (18) is less easily justified on the grounds of the existence of common factors in such VHAR type volatility models.

In order to estimate the parameters of model (15), we resort to a switching algorithm (see details about the estimation technique and its Monte Carlo evaluation in Cubadda, et al., 2017) that is
widely applied in cointegration analysis (see Boswijk and Doornik, 2004, and their references). The strategy consists in
alternating between estimating $\omega$ for a given value of $\beta$ and $\Sigma$, and estimating $\beta$ and $\Sigma$ for a
given value of $\omega$. In details, the procedure goes as follows:

1. Conditional to an (initial) estimate of the $\omega$, estimate $\beta$ and $\Sigma$ by OLS on (15).

2. Premultiplying both the sides of (15) by $\Sigma^{-1/2}$ one obtains

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

(19)

Applying the Vec operator to both sides of the above equation, using the property $\text{Vec}(ABC) = (C' \otimes A)\text{Vec}(B)$, and keeping in mind that the \text{Vec} of a column vector is itself, one gets

$$\Sigma^{-1/2}(Y_{t}^{(day)} - \beta_0) = \left( Y_{t-1}^{(day)} \otimes \Sigma^{-1/2}\beta^{(day)} \right) \text{Vec}(\omega') + \left( Y_{t-1}^{(w)} \otimes \Sigma^{-1/2}\beta^{(w)} \right) \text{Vec}(\omega') + \left( Y_{t-1}^{(m)} \otimes \Sigma^{-1/2}\beta^{(m)} \right) \text{Vec}(\omega') + \Sigma^{-1/2}\epsilon_t,$$

(20)

from which we can finally estimate by OLS the $\omega$ coefficients conditional to the previously obtained estimates of the parameters $\beta$ and $\Sigma$.

3. Switch between steps 1 and 2 till numerical convergence occurs.

As shown by Boswijk (1995), the proposed switching algorithm has the property to increase the Gaussian likelihood in each step.

Note that a numerical stability problem may arise when the number of series is large. As suggested by Cubadda and Guardabschio (2018), a possible solution to this problem is to resort to ridge regression in place of OLS in each of the steps 2-3 above. In particular, ridge estimation can be directly applied to model (20) and to model (15) after having applied the Vec operator to both sides of the related equation, which leads to

$$Y_{t}^{(day)} = \beta_0 + \left( f_{t-1}^{(day)} \otimes I_n \right) \text{Vec}(\beta^{(day)}) + \left( f_{t-1}^{(w)} \otimes I_n \right) \text{Vec}(\beta^{(w)}) + \left( f_{t-1}^{(m)} \otimes I_n \right) \text{Vec}(\beta^{(m)}) + \epsilon_t.$$  

(21)

Clearly, each of the two ridge regressions above requires to fix its own tuning parameter. This can be done following Hoerl \textit{et al.} (1975), namely using the ML estimates to compute the optimal values of the tuning parameters (see Cubadda and Guardabascio, 2018, for further details).

4 Data description

We have considered intra-day data extracted from the NYSE “trade and quote” (TAQ) dataset downloaded from Thomson Reuters. It contains the 250 most liquid assets quoted on New York Stocks Exchange covering the period from 03/01/2006 to 31/12/2014 not including weekends and holidays, for a sample of 2265 trading days. From the dataset, we focus in this study on the asset prices of thirteen major banks. These are, in alphabetical order of the acronyms, (1) BAC:
Bank of America Corporation, (2) BBT: BB&T Corporation, (3) BK: Bank Of New York Mellon Corporation (The), (4) C: Citigroup Inc., (5) COF: Capital One Financial Corporation, (6) JPM: J P Morgan Chase & Co, (7) KEY: KeyCorp, (8) PNC: PNC Financial Services Group, Inc. (The), (9) RF: Regions Financial Corporation, (10) STI: SunTrust Banks, Inc., (11) STT: State Street Corporation, (12): USB U.S. Bancorp and (13) WFC: Wells Fargo & Company. We focus on these assets since it is commonly believed that the banking sector is highly exposed to contagion effects.

Data have been cleaned following the procedure proposed by Barndorff-Nielsen et al. (2009). It consists of the following different steps:

**Steps applied to all data**

P1. Delete entries with a time stamp outside the 9:30 am to 4 p.m. window when the exchange is open.

P2. Delete entries with a bid, ask or transaction price equal to zero.

**Steps applied only to quote data**

Q1. When multiple quotes have the same timestamp, replace all these with a single entry with the median bid and median ask price.

Q2. Delete entries for which the spread is negative.

Q3. Delete entries for which the spread is more than 50 times the median spread on that day.

Q4. Delete entries for which the mid-quote deviated by more than 10 mean absolute deviations from a rolling centered median (excluding the observation under consideration) of 50 observations (25 observations before and 25 after).

**Steps applied only to trade data**

T1. If multiple transactions have the same time stamp: use the median price.

T2. Delete entries with prices that are above the ask plus the bid-ask spread. Similar for entries with prices below the bid minus the bid-ask spread.

T3. Delete entries for which the price deviated by more than 10 mean absolute deviations from a rolling centered median (excluding the observation under consideration) of 50 observations (25 observations before and 25 after).³

³Both Q4 and T3 are very closely related to the procedure by Brownlees and Gallo (2006). Indeed, the median is used in place of the trimmed sample mean, $\bar{p}_i(k)$, and the mean absolute deviation from the median in place of $s_i(k)$.

At the end of the procedure, the number of trades has been reduced from around 225 millions to slightly more than 105 millions. Notice that, although only trade prices are considered in the
analysis, the above cleaning procedure also involves quote data. The reason of this choice is to obtain the more coherent trade prices as possible.

Then, prices have been sampled at 5-minute frequency using the previous point interpolation method and two different realized volatility measures have been computed from the correspondent returns, namely the 5-minute Realized Variance (RV5) already defined in the introduction section and the 5-minute Median Truncated Realized Variance (MedRV5), such as

$$RV5_t \equiv \sum_{j=1}^{M} r_{t,j}^2,$$

where

$$MedRV5_t \equiv \frac{\pi}{6 - 4\sqrt{3} + \pi} \left( \frac{M}{M - 2} \right) \sum_{j=2}^{M-1} \text{med}(|r_{t,j}||r_{t,j-1}||r_{t,j+1}|)^2,$$

where $r_{t,j}$ are the high frequency intra-day returns, observed for $M$ intra-day 5-min periods we have considered each day.

Figures 1 and 2 display the levels as well as the log levels of the 13 realized variance series, over the period of 03/01/2006 to 31/12/2014. We can see later that this distinction has an impact on the interpretation of the factors that we extract from those variables. Indeed, although taking the logs seems natural to get variables with properties closer to the Gaussian distribution, the aggregation of the levels is easier when the goal is to obtain an index from the (weighted) sum of individual volatilities. In Figures 1 and 2, "whole sample" denotes the period 03/01/2006 to 31/12/2014. In the forecasting exercise of Section 4, we will only report forecasting performances for the post-crisis period 03/01/2008 to 31/12/2014 as results might have changed since the financial crisis (Bilio et al., 2012; Ahelegbey et al., 2016). While our study stays valid for highly volatile periods, a model confidence set approach was not able to statistically distinguish the different clusters when such a huge crisis period is included in the sample. This issue, that we can summarize by, all models are bad and cannot be distinguished, has also been noticed by Hecq et al. (2012) for instance. Finally, even though not reported here, ACFs for each series for both the 03/01/2006 to 31/12/2014 and the 03/01/2008 to 31/12/2014 periods show slow decay patterns. Consequently we really have a long-memory feature present in our series and not a spurious phenomenon due to structural breaks (before and after the financial crisis) as in Diebold and Inoue (2001) or Haldrup and Kruse (2014) for instance.

Table 1 illustrates that for the log of the 13 realized volatility series, both the HAR and the FWN models fit pretty well the long-memory feature. We provide OLS estimates of HAR equations (11). We also report the adjusted determination coefficient $\bar{R}^2$ as well as the $p-value$ of the Ljung-Box $\chi^2$ test for the null hypothesis that the first 10 error lags are zero. The goodness of fit, above 0.82

\footnote{Similar patterns emerge for the MedRV.}

\footnote{Similar results are obtained on MedRV. Also we only report the results for the logs of the series while we compare forecasting performances of both levels and log levels in the forecasting section.}
Figure 1: Levels of realized volatilities
Figure 2: Log-levels of realized volatilities
for both the fractional white noise and the HAR is very similar for each asset volatilities. Note that the maximum likelihood estimates of $d$ of the ARFIMA$(0,d,0)$ are rather close to 0.5, one of the example considered by CHL18. Whereas ML bounds the parameter estimate at that value, we have also considered the estimation for the series in first differences before we obtain "unbounded" estimates of $d + 1$. In every case we obtain estimated values for $d + 1$ between 0.5 and 0.54 with a significant difference to 0.5 in only one case at 1% level.

Finally, let us have a look at VAR coefficient matrices in a VAR(1) in order to figure out how close we are to the CHL18 setup. Estimating VARs for the log of the 13 RV5 series as well as for the log of MedRV5 we obtain VAR(5), VAR(1) and VAR(2) with respect to AIC, BIC and HQIC using $p_{\text{max}} = 22$ days. Table 2 provides results for the VAR(1) coefficient matrix chosen by the BIC for the logs of the realized variances. We denote $\hat{A}_{13}$ that estimated matrix for a thirteen dimensional VAR(1) model. We indeed observe as in the model proposed by CHL18 a large value for the diagonal elements and relatively small and often non significant off-diagonal elements.

The hypothesis that the VAR model has a diagonal structure is strongly rejected. Indeed, the null hypothesis that the only non zero coefficients are those of the lags of the dependent variable is overwhelmingly rejected in each equation. These findings suggest that these data may accord with the modeling framework of CHL18. However note that, as discussed in Section 1, the thirteen series might have been generated from a larger VAR, in which the off-diagonal elements are likely to be even smaller. It might well be that some coefficients that are small but not insignificant would have been closer to zero if we had estimated a VAR model with a larger number of volatilities than 13. This would be the case if there is a positive omitted variable bias if one estimates such a smaller system.

5 Forecasting

We compare the forecasting performances of the VHAR and the VHARI with those of univariate models such as the FWN and the HAR. The forecasting exercises are performed using a rolling window of 500 days. Table 3 refers to the levels of MedRV5, whereas Table 4 refers to their logarithmic transformation. In each rolling sample, the FWN is estimated by Gaussian ML, the HAR by OLS, whereas the VHAR and the VHARI have been estimated by the switching algorithm using either OLS or ridge regression. In the latter case, they are denoted respectively as VHAR (R) and VHARI (R) in Tables 3 and 4. For both the VHARI and VHARI (R), the number of indexes $q$ has been determined using the usual information criteria proposed by Schwarz (BIC), Hannan-Quinn (HQIC) and Akaike (AIC). Forecasts are obtained for each sample of 500 observations using the empirical number of factors $\hat{q}$ for that specific sample. What the last four columns of Tables 3 and 4 denote are summary statistics (quartiles and mode) over all rolling windows of 500 days. This gives an overall indication about the parsimony of the different information criteria in choosing $q$.

For all the methods, direct $h$--step ahead forecast for $h = 1, 5, 22$ are produced. As in Cubadda
et al. (2017), the goodness of the forecasts has been evaluated through the average relative mean squared forecast errors (ARMSFE). It is defined as

$$\text{ARMSFE}_{m,h} = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{\text{MSFE}_{m,h,i}}{\text{MSFE}_{HAR,h,i}} \right) \times 100$$  \hspace{1cm} (24)

where $m$ denotes the model (e.g. VHARI or VHAR) and $n$ represents the number of assets, which is equal to 13 in our application. In order to find the set of models which forecast equally well, we rely on the model confidence set analysis by Hansen et al. (2011). In particular, the test for the null hypothesis of equal predictive ability at the 20% level is implemented using a block bootstrap scheme with 5000 resamples.

Looking first at the results regarding the levels of MedRV5 in Table 3, it emerges that univariate models have smaller ARMSFEs than those of the multivariate ones at any time horizons, sometimes by 30%. The good approximation of the FWN model by the HAR is also illustrated in our results since they exhibit similar forecasting performances. However, the differences among the various methods are not always significant. Indeed, for $h = 1, 22$, the set of superior models includes all the models. In the case of the logs, the picture is more clear cut. Table 4 shows that, for $h = 1$, the subset of the best models includes only the univariate ones. There are no significant differences for $h = 22$.

These empirical findings are in sharp contrast with those obtained by Cubadda et al. (2017), which indicate that the VHARI often outperforms the HAR specification in forecasting. However, these authors used a rather different kind of data from those considered here, namely several realized volatility measures of the same equity index for three different markets. Hence, it might well be possible that a common index structure is more appropriate for modelling various measures of the same underlying volatility rather than the realized volatilities of different assets.

Remarkably, the VHAR and the three specifications of the VHARI have similar forecasting performances, thus suggesting that imposing a common index structure is not of particular help for modelling the considered volatilities. Indeed, the best forecasting results for the VHARI are obtained when the number of indexes is selected by the AIC, which of course provides the less parsimonious specification of the model (with the estimated $q$ ranging in median from 6 to 13). Finally, ridge regularization in estimation slightly improves in forecasting when the model is either the VHAR or the VHARI with $q$ selected by the AIC.

Overall, the empirical results seem to indicate that the contagion effect among the various volatilities is to some extent limited or, at least, is not well captured by the two multivariate versions of the HAR model that we have considered here. A scheme, à la CHL18, in which there exists at the multivariate level a large dimensional system with large diagonal autoregressive coefficients in the volatility of the series and tiny but jointly significant off-diagonal elements is not rejected by the data. FWN models as well as HAR, although estimated independently on each variables, seem to keep a footprint of what is referred to as the cross-sectional hidden dependence issue in CHL18.
6 Conclusions

In this paper we have evaluated the performances of various models in forecasting the volatilities of 13 asset returns. Our empirical findings indicate that univariate methods outdo the multivariate ones, although the differences in the associated ARMSFEs are often insignificant according to the model confidence set analysis. Moreover, imposing common components in the VHAR does not significantly improve forecast accuracy.

Overall, this evidence is line with the view that a large dimensional VAR with small contagion effects is the underlying generating mechanism of the considered realized volatilities (Chevillon et al., 2018).

An interesting option is to develop a multivariate model for realized volatilities that, differently from the VHARI, allows for both idiosyncratic and common components. This issue is currently on our research agenda.

References


<table>
<thead>
<tr>
<th>Series</th>
<th>Model</th>
<th>$d$</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\alpha_3$</th>
<th>$\bar{R}^2$</th>
<th>$Q(10)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAC</td>
<td>ARFIMA(0,d,0)</td>
<td>0.497*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.83</td>
<td>0.306</td>
</tr>
<tr>
<td></td>
<td>HAR</td>
<td>–</td>
<td>0.491*</td>
<td>0.263*</td>
<td>0.219*</td>
<td>0.83</td>
<td>0.594</td>
</tr>
<tr>
<td>BBT</td>
<td>ARFIMA(0,d,0)</td>
<td>0.496*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.85</td>
<td>0.049</td>
</tr>
<tr>
<td></td>
<td>HAR</td>
<td>–</td>
<td>0.416*</td>
<td>0.347*</td>
<td>0.215*</td>
<td>0.85</td>
<td>0.108</td>
</tr>
<tr>
<td>BK</td>
<td>ARFIMA(0,d,0)</td>
<td>0.496*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.82</td>
<td>0.109</td>
</tr>
<tr>
<td></td>
<td>HAR</td>
<td>–</td>
<td>0.449*</td>
<td>0.307*</td>
<td>0.214*</td>
<td>0.82</td>
<td>0.103</td>
</tr>
<tr>
<td>C</td>
<td>ARFIMA(0,d,0)</td>
<td>0.497*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.85</td>
<td>0.054</td>
</tr>
<tr>
<td></td>
<td>HAR</td>
<td>–</td>
<td>0.441*</td>
<td>0.339*</td>
<td>0.195*</td>
<td>0.84</td>
<td>0.000</td>
</tr>
<tr>
<td>JPM</td>
<td>ARFIMA(0,d,0)</td>
<td>0.497*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.86</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>HAR</td>
<td>–</td>
<td>0.498*</td>
<td>0.266*</td>
<td>0.205*</td>
<td>0.82</td>
<td>0.782</td>
</tr>
<tr>
<td>KEY</td>
<td>ARFIMA(0,d,0)</td>
<td>0.497*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.85</td>
<td>0.107</td>
</tr>
<tr>
<td></td>
<td>HAR</td>
<td>–</td>
<td>0.444*</td>
<td>0.305*</td>
<td>0.230*</td>
<td>0.86</td>
<td>0.252</td>
</tr>
<tr>
<td>PNC</td>
<td>ARFIMA(0,d,0)</td>
<td>0.496*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.84</td>
<td>0.058</td>
</tr>
<tr>
<td></td>
<td>HAR</td>
<td>–</td>
<td>0.432*</td>
<td>0.342*</td>
<td>0.201*</td>
<td>0.84</td>
<td>0.028</td>
</tr>
<tr>
<td>RF</td>
<td>ARFIMA(0,d,0)</td>
<td>0.497*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.86</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>HAR</td>
<td>–</td>
<td>0.418*</td>
<td>0.356*</td>
<td>0.203*</td>
<td>0.86</td>
<td>0.000</td>
</tr>
<tr>
<td>STI</td>
<td>ARFIMA(0,d,0)</td>
<td>0.496*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.85</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>HAR</td>
<td>–</td>
<td>0.415*</td>
<td>0.353*</td>
<td>0.209*</td>
<td>0.85</td>
<td>0.011</td>
</tr>
<tr>
<td>STT</td>
<td>ARFIMA(0,d,0)</td>
<td>0.490*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.82</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>HAR</td>
<td>–</td>
<td>0.371*</td>
<td>0.384*</td>
<td>0.216*</td>
<td>0.82</td>
<td>0.049</td>
</tr>
<tr>
<td>USB</td>
<td>ARFIMA(0,d,0)</td>
<td>0.497*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.85</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>HAR</td>
<td>–</td>
<td>0.452*</td>
<td>0.343*</td>
<td>0.179*</td>
<td>0.85</td>
<td>0.347</td>
</tr>
<tr>
<td>WFC</td>
<td>ARFIMA(0,d,0)</td>
<td>0.498*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.85</td>
<td>0.006</td>
</tr>
</tbody>
</table>

Note: $d$ is the estimated integration order, $\alpha_i$, for $i = 1, 2, 3$, are the estimated HAR coefficients, $\bar{R}^2$ is the determination coefficient, $p-value$ refers to the Ljung-Box test for the null hypothesis that the first 10 lags are zero, * denotes significance (HCSE) at the 1% significance level.
Table 2: VAR(1) estimated matrix for logs of realized volatilities

\[
\hat{A}_{13} = \\
\begin{pmatrix}
0.52 & -0.04 & -0.02 & 0.01 & -0.06 & 0.02 & -0.01 & -0.07 & 0.12 & 0.06 & -0.04 & -0.01 & -0.05 \\
-0.08 & 0.31 & 0.07 & -0.07 & 0.06 & 0.03 & 0.03 & -0.01 & 0.00 & 0.02 & 0.05 & -0.01 & -0.01 \\
0.03 & 0.07 & 0.39 & -0.03 & 0.03 & 0.11 & -0.01 & -0.01 & -0.04 & 0.00 & 0.12 & 0.00 & 0.00 \\
0.11 & 0.01 & 0.01 & 0.54 & 0.06 & 0.07 & 0.10 & 0.04 & 0.07 & 0.11 & 0.04 & 0.07 & 0.10 \\
0.04 & 0.18 & 0.09 & 0.13 & 0.49 & 0.14 & 0.04 & 0.16 & 0.07 & 0.15 & 0.13 & 0.10 & 0.21 \\
0.05 & 0.01 & 0.18 & -0.00 & 0.01 & 0.38 & -0.10 & 0.00 & -0.10 & -0.07 & 0.06 & 0.00 & 0.02 \\
0.03 & 0.10 & 0.03 & 0.09 & 0.03 & 0.02 & 0.45 & 0.12 & 0.17 & 0.12 & 0.07 & 0.09 & 0.04 \\
-0.05 & 0.00 & -0.04 & -0.04 & 0.08 & -0.06 & 0.02 & 0.30 & -0.05 & -0.02 & 0.05 & 0.07 & 0.00 \\
0.17 & 0.09 & 0.013 & 0.11 & -0.02 & 0.02 & 0.18 & 0.05 & 0.48 & 0.15 & 0.00 & 0.08 & 0.06 \\
0.06 & -0.03 & -0.02 & 0.07 & -0.02 & -0.04 & 0.05 & -0.01 & 0.15 & 0.30 & -0.06 & -0.02 & 0.01 \\
0.00 & 0.03 & 0.14 & 0.02 & 0.04 & 0.03 & 0.05 & 0.08 & -0.03 & -0.03 & 0.34 & 0.06 & 0.06 \\
0.03 & 0.09 & 0.01 & -0.03 & 0.03 & 0.03 & 0.08 & 0.12 & 0.04 & 0.07 & 0.05 & 0.35 & 0.04 \\
-0.00 & 0.08 & 0.02 & 0.11 & 0.18 & 0.09 & 0.04 & 0.12 & 0.00 & 0.04 & 0.08 & 0.12 & 0.43
\end{pmatrix}
\]

Table 3: Forecast comparison for levels of MedRV - post crisis period

<table>
<thead>
<tr>
<th>Method/Criterion</th>
<th>ARMSFE</th>
<th>q</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>h = 1</td>
<td>h = 5</td>
</tr>
<tr>
<td>VHARI/BIC</td>
<td>137.9*</td>
<td>108.3</td>
</tr>
<tr>
<td>VHARI/HQC</td>
<td>136.3*</td>
<td>112.0</td>
</tr>
<tr>
<td>VHARI/AIC</td>
<td>134.6*</td>
<td>111.0</td>
</tr>
<tr>
<td>VHAR</td>
<td>135.5*</td>
<td>118.0</td>
</tr>
<tr>
<td>VHARI (R)/BIC</td>
<td>135.0*</td>
<td>102.4</td>
</tr>
<tr>
<td>VHARI (R)/HQC</td>
<td>131.1*</td>
<td>104.0</td>
</tr>
<tr>
<td>VHARI (R)/AIC</td>
<td>130.3*</td>
<td>110.9</td>
</tr>
<tr>
<td>VHAR (R)</td>
<td>130.3*</td>
<td>110.9</td>
</tr>
<tr>
<td>FWN</td>
<td>97.3*</td>
<td>100.1*</td>
</tr>
<tr>
<td>HAR</td>
<td>100*</td>
<td>100*</td>
</tr>
</tbody>
</table>

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 i-th quartile of the number of factors distribution.
The methods within the superior set of models, as identified by the MCS test, are denoted by *.  

23
<table>
<thead>
<tr>
<th>Method/Criterion</th>
<th>ARMSFE $h = 1$</th>
<th>ARMSFE $h = 5$</th>
<th>ARMSFE $h = 22$</th>
<th>$\hat{q}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>VHARI/BIC</td>
<td>121.6</td>
<td>108.4</td>
<td>111.5*</td>
<td>[1 2 2]</td>
</tr>
<tr>
<td>VHARI/HQIC</td>
<td>117.9</td>
<td>110.1</td>
<td>120.6</td>
<td>[3 3 4]</td>
</tr>
<tr>
<td>VHARI/AIC</td>
<td>117.7</td>
<td>110.8</td>
<td>118.3*</td>
<td>[6 6 7]</td>
</tr>
<tr>
<td>VHAR</td>
<td>118.8</td>
<td>117.2</td>
<td>114.2*</td>
<td></td>
</tr>
<tr>
<td>VHARI (R)/BIC</td>
<td>137.6</td>
<td>115.1</td>
<td>106.9*</td>
<td>[1 1 2]</td>
</tr>
<tr>
<td>VHARI (R)/HQIC</td>
<td>122.7</td>
<td>111.7</td>
<td>112.8*</td>
<td>[2 3 4]</td>
</tr>
<tr>
<td>VHARI (R)/AIC</td>
<td>115.2</td>
<td>111.2</td>
<td>111.4*</td>
<td>[5 8 9]</td>
</tr>
<tr>
<td>VHAR (R)</td>
<td>115.3</td>
<td>111.3</td>
<td>111.5*</td>
<td></td>
</tr>
<tr>
<td>FWN</td>
<td>99.8*</td>
<td>99.2*</td>
<td>101.3*</td>
<td></td>
</tr>
<tr>
<td>HAR</td>
<td>100*</td>
<td>100*</td>
<td>100*</td>
<td></td>
</tr>
</tbody>
</table>

See the notes of Table 3.
RECENT PUBLICATIONS BY CEIS Tor Vergata

Wavelet analysis for temporal disaggregation
Chiara Perricone
*CEIS Research Paper, 444*, October 2018

Ambiguous economic news and heterogeneity: What explains asymmetric consumption responses?
Luisa Corrado, Robert Waldmann, Donghoon Yoo
*CEIS Research Paper, 443*, August 2018

New Technologies and Costs
Vincenzo Atella and Joanna Kopinska
*CEIS Research Paper, 442*, August 2018

On Competing Mechanisms under Exclusive Competition
Andrea Attar, Eloisa Campioni and Gwenaël Piaser
*CEIS Research Paper, 441*, August 2018

About Local Projection Impulse Responde Function Reliability
Luca Brugnolini
*CEIS Research Paper, 440*, August 2018

Estimating Models with Dynamic Network Interactions and Unobserved Heterogeneity
Luisa Corrado and Salvatore Di Novo
*CEIS Research Paper, 439*, August 2018

Political Turnover and the Performance of Local Public Enterprises
Andrea De Meo and Lorenzo Ferrari
*CEIS Research Paper, 438*, August 2018

Home Sweet Home: the Effect of Sugar Protectionism on Emigration in Italy, 1876-1913
Carlo Ciccarelli, Alberto Dalmazzo and Daniela Vuri
*CEIS Research Paper, 437*, June 2018

Fiscal Compact and Debt Consolidation Dynamics
Luca Brugnolini and Luisa Corrado
*CEIS Research Paper, 436*, June 2018

Protection for Sale with Price Interactions and Incomplete Pass-Through
Barbara Annicchiarico and Enrico Marvasi
*CEIS Research Paper, 435*, June 2018