

A forecast based comparison of restricted realized covariance models*

Matteo Bonato

Swiss Finance Institute and University of Zurich

Massimiliano Caporin

Department of Economics and Management “Marco Fanno”

University of Padova

Angelo Ranaldo

Swiss National Bank

January 13, 2010

*The paper was written while the second author was visiting the Chiang Mai University, Faculty of Economics, Thailand, whose hospitality is gratefully acknowledged.

-The views expressed herein are those of the authors and not necessarily those of the Swiss National Bank, which does not accept any responsibility for the contents and opinions expressed in this paper.

Abstract: Models for realized covariance matrices may suffer for the curse of dimensionality as more traditional multivariate volatility models (such as GARCH and stochastic volatility). Within the class of realized covariance models we focus on the Wishart specification introduced by Gouriéroux et al. (2009) and analyse here the forecasting performances of the parametric restrictions discussed in Bonato et al. (2009) which are motivated by asset features. Our purpose is to verify if restricted model forecasts are statistically equivalent to full model specification, a result that would support the use of restrictions when the problem cross sectional dimension is large.

Keywords: Realized covariance, WAR, HAR, multivariate volatility forecasts.

JEL codes: C32, C53, C52.

1 Introduction

Since mid 90's the availability of high frequency data attracted considerable interest in the econometric, statistical and financial literature. On the one side, many authors considered the filtering problems associated with these data, their use for market microstructure studies, or their direct modelling, see Dacorogna et al. (2001) for a survey. On the other side, high frequency data gave rise to a relevant research area, that of realized volatility, started by the seminal contributions of Andersen et al. (2000, 2001). This strand of the financial econometrics literature could be further divided into two subsets, the first including the theoretical studies dealing with the appropriate measurement of the realized volatility, the second, more empirical, tackling the problem of modeling realized volatility sequences for their financial applications, see the paper by McAleer and Medeiros (2008) for a survey.

The present paper belongs to the last group and focuses on a specific set of models, the Wishart Auto Regressive (or WAR) introduced by Gouriéroux et al. (2009), and then analyzed by Bonato et al. (2009). Realized covariance models suffer, as well as many other covariance models, of the so-called curse of dimensionality, that is, the parameter number increases at an order larger than the model cross-sectional dimension. For this reason, restricted parameterizations have been proposed in many frameworks, see Bauwens et al. (2006), Silvennoinen and Terasvirta (2009), Caporin and Paruolo (2009), among others. Within the WAR framework, a set of possible restrictions have been introduced by Bonato et al. (2009) and motivated by economic and financial features of the analyzed variables. As an example, we may impose that the realized covariance dynamic is constrained to be equal across assets

belonging to the same economic sector. When the cross-sectional dimension is small, full models can be still estimated, and compared to restricted specifications in order to verify what are their effect on model performances. In this paper we present an empirical comparison of full and restricted WAR models whose purpose is to verify how close are restricted specifications to the most general model.

We consider a dataset of 12 large cap US stocks sampled at a 5 minute frequency, which are used to determine the daily realized covariance matrices in the range January 2003 to December 2008. Over these series we fit a set of alternative WAR specifications and compare them in term of forecast performances by using statistical criteria. In details, we consider the Amisano and Giacomini (2006) Weighted Likelihood Ratio test, the pairwise model comparison based on two multivariate loss functions similar to those discussed in Patton and Sheppard (2009) and Clements et al. (2009), and the Model Confidence Set of Hansen et al. (2005).

Our results confirm those obtained by previous studies, suggesting that the choice of the approach used for the comparison could lead to different ranking of alternative models. Nevertheless, we provide evidence showing that restricted model specification have often forecasting performances better than fully specified models.

The paper proceeds as follow. Section 2 describes the model we consider, the parametric restrictions driven by economic features of the data, the parameter interpretation and the model estimation. Section 3 deals with the methods for the forecast evaluation and comparison across multivariate volatility models while Section 4 presents the empirical results. Section 5

concludes.

2 Modelling variance spillovers

We now introduce the model used to analyze and forecast the sequence of realized variance/covariance matrix of the 12 assets in our study. We then describe the set of alternative parametric restrictions to the spillovers between variances that help to reduce the complexity of the model estimation, and might be motivated by some economic criterion (e.g. assets classification) or data driven. As we mention in the introduction, we presume that realized covariance sequences are available and we will not tackle the problem of the optimal estimation of realized variances and covariances.

2.1 The WAR model

Once we have computed the series of realized variance/covariance matrices from the intra-day returns, it is of fundamental importance for our purpose to adopt a model which is feasible to estimate even with a large set of assets, that guarantees the positive definitiveness of the forecasted covariance matrix, and whose coefficients keep their interpretability. This last aspect will be even more important if the quality of covariance matrix forecasts would be measured according to an economic criterion (such as the returns on various optimized portfolio).

A model that satisfies the previously itemized requirement is the realized Wishart Autoregressive model (WAR) of [Gourieroux et al. \(2009\)](#). As shown in [Bonato et al. \(2009\)](#), this model is particularly suitable to test for variance

spillover between assets and in their paper the authors propose four specifications to restrict the interaction between past variance/covariances with their contemporaneous values.

Denote by Y_t the time t (realized) covariance for a group of n assets. The conditional mean equation given by the WAR(1) model reads (see the Appendix for the definition and properties of the WAR model, and [Gourieroux et al. \(2009\)](#) for additional details on WAR)

$$Y_{t+1} = MY_tM' + K\Sigma + \eta_{t+1}. \quad (1)$$

where η_{t+1} is a zero-mean innovation term with a non-standard density.

This simple formulation resembles in many aspects the classical AR(1) conditional mean widely used in the literature to model and forecast realized volatility. The WAR(1) model states that the covariance matrix we observe today depends directly on its value on the previous day. The parameter matrix Σ and the scalar K arise from the definition of the WAR(1) model. In particular, the value of this latter parameter, named the degrees of freedom, is fundamental to assess the appropriateness of the model. To ensure positive definiteness of the fitted realized covariances the degree of freedom must satisfy $K > n - 1$. See for example [Bonato \(2009\)](#) for an in-depth discussion on the topic.

Of course, more than one lag can be used in the model, leading to the WAR(p) model, $p \geq 1$, which reads

$$Y_{t+1} = \sum_{j=1}^p M_j Y_{t+1-j} M_j' + K\Sigma + \eta_{t+1}. \quad (2)$$

As discussed in [Corsi \(2009\)](#), agents on the financial markets operate

with different time horizons (they are heterogeneous). Such a behavior could generate long-memory patterns. To capture this feature Corsi (2009) introduces a new univariate realized volatility model named Heterogeneous Auto Regression, HAR, that capture the different agents horizons, by making today realized volatility a function of daily, weekly and monthly past realized volatility component. Along with the classical WAR model and its restricted specifications described below, we will also test the economic values of the restriction on the HAR-WAR model proposed in Bonato et al. (2009). Define the k -period realized covariance matrix component by the average of the single-period realized covariance matrices:

$$Y_{t-k:t-1} = \frac{1}{k} \sum_{j=1}^k Y_{t-j} \quad (3)$$

Combining a WAR(p) structure with the temporal aggregation induced by the HAR model, we write the process Y_t as

$$Y_t = M_1 Y_{t-1} M_1' + M_2 Y_{t-5:t-1} M_2' + M_3 Y_{t-22:t-1} M_3' + K \Sigma + \eta_{t+1}. \quad (4)$$

2.2 Interpretation of the coefficients

As previously mentioned, one of the advantages of the WAR is that the interpretation of the coefficients in the model is not lost. Bonato et al. (2009) propose to restrict the parametric space setting some parameters to zero according to economic criteria. Such a strategy helps solving the so called ‘curse of dimensionality’, that is, the number of parameters is a power function of the cross-sectional model dimension. The main contribution of this paper is to check whether imposing these restrictions affects the forecasting

performances of the model. Equivalently, by working with a cross-sectional dimension allowing the estimation of a fully parameterized model, we verify if the forecasting performances of a restricted specification are statistically different from those of the fully parameterized model.

In this section we briefly introduces the four different restricted parameterizations proposed in Bonato et al. (2009) and discuss their economic interpretation.

Consider the simple WAR(1) model as in Eq. (A.4):

$$Y_{t+1} = MY_tM' + K\Sigma + \eta_{t+1}. \quad (5)$$

Assume that our portfolio consists of n stocks classified into N groups according to some economic (or data-driven) criterion (for instance the economic sector, the company market dimension, or or the existence of common patterns in realized variances and covariances).

The N groups have dimension n_i with $\sum_i^N n_i = n$. In addition, the assets are ordered following a group rule, that is, assets from 1 to n_1 belong to group 1, assets from $n_1 + 1$ to $n_1 + n_2$ belongs to group 2, and so on. Given this asset classification, the autoregressive matrix M may be partitioned as follows:

$$M = \begin{pmatrix} M_{11} & \cdots & M_{1N} \\ \vdots & M_{ii} & \vdots \\ M_{N1} & \cdots & M_{NN} \end{pmatrix}, \quad (6)$$

where M_{ij} is a matrix of dimension $n_i \times n_j$.

Bonato et al. (2009) propose the following specifications:

- (i) $M_{ij} = \mathbf{0} \quad \forall i \neq j, \quad i, j = 1, \dots, N,$

$$(ii) \quad M_{ij} = \mathbf{0} \text{ and } M_{ii} = \alpha_i(\mathbf{1}_{n_i}\mathbf{1}'_{n_i}), \quad \forall i \neq j, \quad i, j = 1, \dots, N$$

$$(iii) \quad M_{ij} = \mathbf{0} \text{ and } M_{ii} = (\alpha_{i,1}, \dots, \alpha^{i,n_i})(\mathbf{I}_{n_i}), \quad \forall i \neq j, \quad i, j = 1, \dots, N$$

$$(iv) \quad M_{ij} = \mathbf{0} \text{ and } M_{ii} = \alpha_i(\mathbf{I}_{n_i}), \quad \forall i \neq j, \quad i, j = 1, \dots, N$$

where $\mathbf{1}_{n_i}$ is a $n_i \times 1$ vector of ones and \mathbf{I}_{n_i} is the identity matrix of dimension n_i .

If assets belonging to the same group share common reactions to shocks, we can hypothesize, to some extent, that their co-volatilities also have a similar behavior. If the groups are sector-specific, model (i) implies that the variances and covariances of each asset are only influenced by the variances and covariances of assets belonging to the same class. Therefore, no volatility spillover exists between assets belonging to different sectors. This model is called *block WAR*. The number of parameters that needs to be estimated is $n(n+1)/2 + \sum_{i=1}^N n_i^2$, along with the degrees of freedom K .

A further reduction of the number of parameters is obtained by imposing a single parameter for each group, as shown in model (ii). In this case, the variance and covariance of each asset belonging to, say, group j depends on the past values of itself, on the past values of the variances of the other assets of the same group and on the covariances with those assets via a function of the unique parameter α_j . We refer to this model as the *restricted block WAR*. This models contains $n(n+1)/2 + N$ parameters in M and Σ plus K .

Model (iii) relaxes the assumption of spillover between assets belonging to the same sector. It assumes each matrix M_{ii} , $i = 1, \dots, N_i$ to be diagonal, i.e. the autoregressive matrix M is diagonal. In this case grouping the assets according to some criterion does not affect the parametric space. This model

is named *diagonal WAR*. For this model, n parameters need to be estimated in the matrix M , plus the $n(n + 1)/2$ parameters in Σ and the degrees of freedom K . One of the implications of the diagonal structure for M is that each realized variance is only a function of its past values.

If one assumes again that assets belonging to the same sector have common dynamics for the variance, or she can find a way to group assets whose volatilities obey the same process, the number of parameters can be further reduced. This is the case for model (iv). For each group a single parameter is taken to model the dynamics of the variances for the assets in the considered group, i.e. the elements on the diagonal of each M_{ii} , $i = 1, \dots, N$ are all equal. In total only $N + n(n + 1)/2 + 1$ parameters are required in this model. We refer to this model as the *restricted diagonal WAR*.

It is worth mentioning that the specifications (i)-(iv) are only a subset of all the possible specifications of the WAR model. In fact, we set all the off-diagonal blocks to zero. The assumption $M_{ij} = \mathbf{0} \forall i \neq j, i, j = 1, \dots, N$ can be replaced by the same structure we imposed on the matrices M_{ii} : full, scalar, diagonal and restricted diagonal. This allows us to consider not only the interactions between assets belonging to the same group, but also interactions between a limited set of groups.

2.3 Model estimation

Under the assumption that $K > n - 1$ Gouriéroux et al. (2009) show that:

- i) K and Σ are identifiable while the autoregressive coefficients in M (and thus M_1, M_2 and M_3) are identifiable up to their sign.

ii) Σ is first-order identifiable up to a scale factor and M is first-order identifiable up to its sign. The degree of freedom K is not first-order identifiable but is second-order identifiable.

Following Gouriéroux et al. (2009), the first-order conditional moments can be used to calibrate the parameters in M and Σ , up to the sign and scale factor, respectively.

We estimate the parameters using non-linear least squares:

$$\left(\hat{M}, \hat{\Sigma}^*\right) = \text{Argmin}_{M, \Sigma^*} S^2(M, \Sigma^*) \quad (7)$$

where

$$\begin{aligned} S^2(M, \Sigma^*) &= \sum_{t=2}^T \sum_{i < j} \left(Y_{ij,t} - \sum_{k=1}^n \sum_{l=1}^n Y_{kl,t-1} m_{ik} m_{lk} - \sigma_{ij}^* \right)^2 \\ &= \sum_{t=2}^T \| \text{vech}(Y_t) - \text{vech}(MY_{t-1}M' + \Sigma^*) \|^2 \end{aligned} \quad (8)$$

and $\Sigma^* = K\Sigma$.

To estimate the degrees of freedom we follow the strategy proposed in Bonato (2009) which has been shown to be less sensible to the presence of extreme events in the co-volatility process. Consider a portfolio allocation $\alpha \in \mathbb{R}^n$. We know that the unconditional distribution of Y_t is a $W(K, 0, \Sigma(\infty))$, a centered Wishart distribution. We can therefore easily show that

$$\alpha'Y_t\alpha \sim \text{Ga}\left(\frac{K}{2}, 2\alpha'\Sigma(\infty)\alpha\right), \quad (9)$$

i.e. the distribution of the portfolio with allocation α is a gamma distribution with the degrees of freedom K as shape parameter, (see also Meucci, 2005). An unbiased estimator of K can be obtained simply via maximum likelihood by fitting a gamma distribution to the process $\alpha'Y_t\alpha$.

3 Comparing realized covariance models

We previously described the various parameter restriction design we could use for WAR and HAR-WAR models. The primary purpose of our work is to compare the forecasts of the realized covariance matrix provided by alternative restricted parameterizations and to compare them to the forecasts of the full model. We denote the alternative specifications by a subscript $i = 1, 2, \dots, m$, and the forecast of time t covariance by $\Sigma_{t,i}^f$. In the following, even if we do not explicitly specify it, all forecasts are conditional to the information set at time $t - 1$. We assume that for all specifications a set of one-step-ahead forecasts are available in a specific range from time $T + 1$ to $T + W$. Furthermore, the observed time t returns of the assets are included in the vector R_t , and the models are estimated with a rolling window approach.

To compare the models we first consider the Weighted Likelihood Ratio test of Amisano and Giacomini (2006). The test compares the out-of-sample forecast performances of two models by resorting to a weighed likelihood function comparison. Within this paper we use the normal likelihood based on the realized covariance forecasts and the observed returns

$$f_{i,t} = -\frac{1}{2} \ln \left| \Sigma_{t,i}^f \right| - \frac{1}{2} R_t' \left(\Sigma_{t,i}^f \right)^{-1} R_t. \quad (10)$$

Given two alternative models i and j and following Amisano and Giacomini (2006), we first compute the quantities

$$WLR_{t,ij} = w(R_t) (f_{i,t} - f_{j,t}) \quad (11)$$

where $w(R_t)$ is a weighting function which could be designed in order to focus on specific regions of the returns density. In our setup we set the weights

to 1 without focusing on specific areas of the forecast density, such as the tails. For additional details see Amisano and Giacomini (2006). The null hypothesis of equal predicting ability of the two models i and j is tested by using the statistic

$$LR_{ij} = \frac{\overline{WLR}_{ij}}{\sqrt{Var(\overline{WLR}_{ij})}} \quad (12)$$

where $\overline{WLR}_{ij} = \frac{1}{W} \sum_{l=1}^W WLR_{T+l,ij}$ and $Var(\overline{WLR}_{ij})$ is obtained by a heteroskedasticity and autocorrelation consistent (HAC) estimator. Under the null hypothesis, LR_{ij} is asymptotically distributed as a standardized normal and, if the null is rejected, its sign could be used to determine the preferred model.

The second approach we consider for the comparison of alternative WAR specifications considers two loss functions for multivariate volatility models. Following Patton and Sheppard (2009), and Clements et al. (2009), we define the two following loss functions

$$L_{i,t}^1 = \frac{1}{k^2} vec(\Sigma_{t,i}^f - \Sigma_t)' vec(\Sigma_{t,i}^f - \Sigma_t), \quad (13)$$

$$L_{i,t}^2 = trace\left(\left(\Sigma_{t,i}^f\right)^{-1} \Sigma_t\right) - \log\left(\left|\left(\Sigma_{t,i}^f\right)^{-1} \Sigma_t\right|\right). \quad (14)$$

which are derived from Patton and Sheppard (2009). The function $L_{i,t}^1$ is a multivariate mean squared error function, while $L_{i,t}^2$ is included in the class of robust loss functions defined in Patton and Sheppard (2009) and similar to a quasi likelihood loss function.

Then, to verify the null that $E[L_{i,t}^w] = E[L_{j,t}^w]$, $w = 1, 2$ we can use a Diebold-Mariano-type test (see Diebold and Mariano, 1995, and West, 1996

and 2000), computing the test statistic

$$\begin{aligned} d_{ij,t}^w &= L_{i,t}^w - L_{j,t}^w \\ L_{ij}^w &= \frac{\bar{d}_{ij}^w}{\sqrt{\text{Var}(\bar{d}_{ij}^w)}} \end{aligned} \quad (15)$$

where $\bar{d}_{ij}^w = \frac{1}{W} \sum_{l=1}^W d_{ij,T+l}^w$, and $\text{Var}(\bar{d}_{ij}^w)$ is obtained by a HAC estimator.

Both the Amisano-Giacomini test and the comparison of models by Diebold-Mariano tests consider two models only. However, when many alternative models must be considered, other methods are needed. Some popular approaches are the Reality Check of White (2000), the Superior Predictive Ability test of Hansen (2005), and the Model Confidence Set (MCS) of Hansen et al. (2005). We consider here the last method, which allows creating a set of model whose forecasting performances are statistically equivalent. The MCS uses as inputs all pairwise loss differentials $d_{ij,t}^w$ for a given loss function w and for all $i, j = 1, 2, \dots, P, i \neq j$, where P is the total number of fitted models. The MCS then proceeds by performing a sequential elimination procedure testing on a set of models \mathcal{M}_l the following null hypothesis $H_0 : E[\bar{d}_{ij,t}^w] = 0$, with $i > j$ and for all $i, j \in \mathcal{M}_l$. The initial set \mathcal{M}_1 contains all models, and if the null hypothesis is rejected the worst performing model is excluded from the set. The procedure works then iteratively until the null is not rejected. Each step performs thus two operations at a generic iteration l : verify the null hypothesis and stop if accepted; if the null is rejected, identify the worst performing model and remove it from the set. To verify the null hypothesis Hansen et al. (2005) proposed two different test statistics based on the quantity (15) but where the variance is computed from a bootstrap procedure (see

Hansen et al., 2005 for details). The two statistics are the following

$$T_R = \max_{i,j \in \mathcal{M}_t} |L_{ij}^w|, \quad (16)$$

$$T_{SQ} = \sum_{i,j \in \mathcal{M}, i > j} (L_{ij}^w)^2. \quad (17)$$

Given that the covariances across the forecasts produced by the models included in a specific set are not null, the test statistics have non-standard and complicated distribution. To determine the p-values of the test statistic a bootstrap approach has been proposed by Hansen et al. (2005). Given a specific confidence level α the null hypothesis can thus be verified by determining the bootstrapped p-values. If it is rejected, the worst performing model is identified as

$$i = \arg \max_{i \in \mathcal{M}} \sum_{j \in \mathcal{M}} \bar{d}_{ij}^w \left(\text{Var} \left(\sum_{j \in \mathcal{M}} \bar{d}_{ij}^w \right) \right)^{-1}.$$

The MCS method was originally proposed for the comparison of univariate volatility forecasts but Patton and Sheppard (2009) suggest it could be of interest also in the multivariate framework, a claim supported by the analysis in Clements et al. (2009). In this paper the MCS will be used as a tool for comparing nested models with respect to their forecasting performances to verify the null hypothesis that a restricted model provides forecast statistically equivalent to those produced by an unrestricted model.

4 Empirical analysis

4.1 Dataset description

To test whether there is economics value in imposing restriction to the variance spillovers between assets, we used a dataset consisting of 12 stocks quoted at NYSE. These are: Boeing Co., Bank of America, Citigroup Inc., Caterpillar Inc., Fedex, Hewlett Packard Co., IBM, JP Morgan Chase Co., Kraft Food Inc., Procter Gamble Co., Time Warner Inc. and Texas Instrument Inc. We then grouped them in 4 assets class according to their sector: Financial, Capital Goods, Technology and Consumer Service. See Table 2 for the sector specific clustering of the dataset.

The dataset consists of one-minute intra-day prices over the period going from January 2, 2003 to December 31, 2008. The data were provided by tickdata.com. After the deletion of holidays and days with no match across the sample, we were left with 1511 trading days. To construct the series of realized covariance matrices we adopted the classical estimator presented in Andersen et al. (2003) and Barndorff-Nielsen and Sheühard (2004) and used, for example, in De Pooter et al. (2006):

$$Y_t = \sum_{i=1}^I r_{t-1+ih,h} r'_{t-1+ih,h}, \quad (18)$$

where Y_t denotes the realized covariance matrix at time t . $r_{t-1+ih,h} \equiv p_{t-1+ih} - p_{t-1+(i-1)h}$ denotes the $(n \times 1)$ vector of returns for the i -th intra-day period on day t , for $i = 1, \dots, I$, and with $n = 12$ the number of assets. I is the number of intra-day intervals, each of length $h \equiv 1/I$. In our case, with a frequency of one minute, $I = 390$.

4.2 Forecast based comparisons

The forecasting period spans the interval from January 2, 2005 until the end of the sample. Using a rolling window of 100 days to estimate the parameters of the model, we computed 755 one-day-ahead forecasts of the realized variance/covariance matrix. We repeated this operation for all the 10 models as in Table 1 and stored the results to compute the test presented in Section 3: the Amisano-Giacomini and the Diebold-Mariano tests for pairwise model comparison, and the Model Confidence Set method for determining the set of models providing statistically equivalent forecasts.

We start by analyzing the outcomes of the Amisano-Giacomini test as reported in Tables 3 and 4. These Tables contain the test statistic and the corresponding p-values, respectively. Several rejections of the null of equal predicting ability are detected, and the sign of the test statistics is used to identify the preferred models: positive values indicates a preference for row models, while negative values a preference for column models. We can identify a subset of models which are preferred to a number of alternatives. These models are the Restricted Block Diagonal WAR (M3), the Block Diagonal HAR-WAR (M7), and the Restricted Block Diagonal HAR-WAR (M8). Notably, the full models (both HAR-WAR and WAR) are inferior to several restricted specifications. This result suggests that restricted specifications could be as good as, if not superior, to fully parameterized specifications.

Tables 5 to 8 report the Diebold-Mariano test statistic and the corresponding p-values for the loss functions defined in (13) and (14), respectively. If we consider the mean squared error loss function (13), Tables 5 and 6, we note that all models are statistically equivalent at the 5% level, while some

differences are detected only at the 10% level. Differently, if we move to the robust loss function (14) results are sensibly different (see Tables 7 and 8). In fact, at the 1% confidence level, fully parameterized models (M1 and M6) are generally statistically preferred to their corresponding restricted specifications (full WAR is preferred to restricted WAR, and full HAR-WAR is preferred to restricted HAR-WAR). There is however a common and interesting exception, restricted diagonal models (M5 and M10) are equivalent to fully parameterized specifications. The restricted diagonal models are also preferred to other restricted specifications. Across classes (WAR and HAR-WAR), we note some common preference for the restricted diagonal HAR-WAR (M10) with respect to many other models.

Finally, Table 9 reports the results of the MCS approach for the two loss functions considered, the mean squared error and the quasi likelihood loss functions. Both tables report, for a given loss function, the MCS results for the two test statistics proposed by Hansen et al. (2005). In turn, results are composed by a ranking of models and the corresponding p-values. Focusing on Table 9 (right columns), first and second columns, mean squared error loss function (13), test statistic in (16), we observe that for confidence levels higher than 20% all models are equivalent.

Moving to the loss function in (14) results are different (see Table 9 left columns). In fact, at the 10% confidence level only a single model is included in the confidence set, model M8, the Restricted Block-Diagonal HAR-WAR. Interestingly, this is also the model with the highest ranking in Table 9 right columns. Lowering the confidence level to 5%, the optimal model set contains also models, M2, M9, M3, and M7, all restricted models. Again, the same models

have high rankings according to Loss Function 1. Finally, we observe that the rankings and the p-values have similar patterns with respect to the MCS test statistics.

Summarizing, results are partially influenced by the choice of the loss functions and the approach used. However, our study shows that restricted models have performances similar, if not superiors, to fully parameterized models. This outcome, is common over different approaches for direct model comparison.

5 Conclusion

This paper presents an empirical comparison of alternative Wishart Auto Regressive models differing in their parameterizations. We evaluate the models within a moderate cross-sectional dimension, 12 assets, and by using direct forecast evaluation approaches. The tests we consider provide some evidence on the possible preference of restricted specifications over fully parameterized models. The results suggests the need of further investigations based on indirect model evaluation, in particular within a portfolio allocation or risk management perspectives. These extensions are left to future researches.

References

- [1] Amisano, G., and Giacomini, R., 2007, Comparing density forecasts via weighted likelihood ratio tests, *Journal of Business and Economic Statistics*, 25-2, 177-190.

- [2] Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H., 2001, The distribution of realized stock return volatility, *Journal of Financial Economics*, 61-1, 43-76.
- [3] Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P., 2000, Exchange rate return standardized by realized volatility are (nearly) Gaussian, *Multinational Finance Journal* 4-3/4, 159-179.
- [4] Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P., 2001, The distribution of exchange rate volatility, *Journal of the American Statistical Association* 96-453, 42-55.
- [5] Asai, M., M. McAleer and J. Yu (2006), Multivariate stochastic volatility: A review, *Econometric Reviews*, 25, 145-175.
- [6] Barndorff-Nielsen O.E. and Shephard N. Econometric analysis of realized covariation: high-frequency based covariance, regressions, and correlation in financial economics, *Econometrica*, 2004, 72 (3), 885-925
- [7] Bauwens, L., Laurent, S., and Rombouts, J.V.K., 2006, Multivariate GARCH models: A survey, *Journal of Applied Econometrics*, 21, 79–109.
- [8] Billio, M., Caporin, M. and Gobbo, M., 2006, Flexible dynamic conditional correlation multivariate GARCH for asset allocation, *Applied Financial Economics Letters*, 2, 123-130.
- [9] Bonato, M., Caporin, M., and Ranaldo, A., 2009, Forecasting realized covariances with a Block structure VAR model, *Swiss National Bank Working Paper* 2009-03.

- [10] Bonato M., Estimating the degrees of freedom of the realized volatility Wishart autoregressive model, 2009, manuscript available at <http://ssrn.com/abstract=1357044>
- [11] Caporin, M., and Paruolo, P., 2009, Structured multivariate volatility models, Department of Economics “Marco Fanno”, University of Padova, Working Paper 90.
- [12] Chib, S., Omori, Y., and Asai, M., 2009, Multivariate Stochastic Volatility, in Andersen, T.G., Davis, R.A., Kreiß, J.P., and Mikosch, T., Handbook of Financial Time Series, Springer.
- [13] Clements, A., Doolan, M., Hurn, S., and Becker, M., 2009, On the efficacy of techniques for evaluating multivariate volatility forecasts, NCER working paper series.
- [14] Corsi, F., 2009, A Simple Approximate Long-Memory Model of Realized Volatility, *Journal of Financial Econometrics*, 7 (2), 174-196.
- [15] Dacorogna, M., Gencay, R., Muller, U.A., Olsen, R., and Pictet, O., 2001, An introduction to high-frequency finance, Academic Press.
- [16] De Pooter, M., M. Martens and D. van Dijk (2008), Predicting the daily covariance matrix for S&P 100 stocks using intraday data - but which frequency to use?, *Econometric Reviews* 27, 199-229
- [17] Diebold, F.X., Mariano, R.S., 1995, Comparing predictive accuracy, *Journal of Business and Economic Statistics*, 13-3, 253-263.

- [18] Engle, R.F., and Colacito, R., 2006, Testing and valuing dynamic correlations for asset allocation, *Journal of Business and Economic Statistics*, 24, 238-253.
- [19] Engle, R.F., and Sheppard, K., 2008, Evaluating the specification of covariance models for large portfolios, available at www.kevinsheppard.net.
- [20] Ferreira, M.A. and Lopez, J.A., 2005, Evaluating interest rate covariance models within a value-at-risk framework, *Journal of Financial Econometrics*, 3, 126-168.
- [21] Gouriéroux, C., Jasiak, J., and Sufana, R., 2009, The Wishart autoregressive process of multivariate stochastic volatility, *Journal of Econometrics*, 150-2, 167-181
- [22] Hansen, P.R., 2005, A test for superior predictive ability, *Journal of Business and Economic Statistics*, 23-4, 365-380.
- [23] Hansen, P.R., Lunde, A., Nason, J.M., 2005, Model confidence sets for forecasting models, Federal Reserve Bank of Atlanta Working Paper 2005-7.
- [24] McAleer M. and M. Medeiros (2008), Realized volatility: a review, *Econometric Reviews*, 27-(1-3), 10-45.
- [25] Meucci, A., 2005, *Risk and Asset Allocation*, Springer

- [26] Mincer, J., and Zarnowitz, V., 1969, The evaluation of economic forecasts. In: Mincer J (ed) Economic Forecasts and Expectations, Columbia University Press.
- [27] Muirhead R.J. 1982, Aspects of multivariate statistical theory, Wiley Series in Probability and Mathematical Statistics
- [28] Patton, A.J., and Sheppard, K., 2009, Evaluating volatility and correlation forecasts, in Andersen, T.G., Davis, R.A., Kreiß, J.P., and Mikosch, T., Handbook of Financial Time Series, Springer.
- [29] Silvennoinen, A., and Terasvirta, T., 2009, Multivariate GARCH models, in Andersen, T.G., Davis, R.A., Kreiß, J.P., and Mikosch, T., Handbook of Financial Time Series, Springer.
- [30] West, K.D., 1996, Asymptotic inference about predictive ability, *Econometrica*, 64-5, 1067-1084.
- [31] West, K.D., 2006, Forecast evaluation, In: Elliott G, Granger C, Timmermann A (eds) Handbook of Economic Forecasting, North Holland Press, Amsterdam.
- [32] White, H., 2000, A reality check for data snooping, *Econometrica*, 68-5, 1097-1126.

A Appendix

A.1 The Wishart Autoregressive model

Denote by Y_t the time t (realized) covariance for a group of n assets. The sequence of stochastic positive definite Y_t matrices is said to follow a Wishart process if the following relations hold. At first, the (realized) covariance may be represented as a sum of underlying stochastic processes

$$Y_t = \sum_{k=1}^K x_{k,t} x_{k,t}' \quad (\text{A.1})$$

where $x_{k,t}$, $k = 1, 2, \dots, K$ are independent Gaussian VAR(1) processes of dimension n with a common autoregressive parameter matrix M and common innovation variance Σ :

$$x_{k,t} = Mx_{k,t-1} + \epsilon_{k,t}, \quad \epsilon_{k,t} \stackrel{i.i.d.}{\sim} N(0, \Sigma). \quad (\text{A.2})$$

When Y_t is defined as in (A.1) and (A.2) we say it follows a WAR process of order 1, denoted $W[K, M, \Sigma]$. The transition density of WAR(1) depends on the following parameters: K , the scalar degree of freedom (the number of underlying VAR processes), strictly greater than $n - 1$ (the number of assets minus one); M , the $n \times n$ matrix of autoregressive parameters; and Σ , the $n \times n$ symmetric and positive definite matrix of innovation covariances. An important property of the Wishart distribution is that the matrices Y_t are positive definite if and only if $K \geq n$ and for a non-centered Wishart specification, the distribution of Y_t possesses a density function only when $K > n - 1$ (hence the condition above). Thus, for $K < n - 1$ no density can be defined and for $K < n$ the process Y_t is given by a sequence of singular

covariance matrices with degenerate Wishart distribution (Muirhead, 1982). We stress that the interpretation of Y_t from latent Gaussian VAR(1) processes is valid for integer valued K only and, in general, any economic or financial interpretation of the latent processes $(x_{k,t})$ is not necessary.

From Proposition 2 in Gouriwroux et al. (2009) we have:

$$E_t(Y_{t+1}) = MY_tM' + K\Sigma. \quad (\text{A.3})$$

The first conditional moment is thus an affine function of the lagged values of the volatility process. In particular, the VAR(1) process is a weak linear AR(1) process. More precisely we get:

$$Y_{t+1} = MY_tM' + K\Sigma + \eta_{t+1}, \quad (\text{A.4})$$

where η_{t+1} is a matrix of stochastic errors with a zero conditional mean. The error term η is a weak white noise, since it features conditional heteroskedasticity and, even after conditional standardization, is not identically distributed.

In general, VAR processes with higher autoregressive order p may be considered and the Wishart process can be easily extended to include more autoregressive lags. This is accomplished by replacing the conditioning matrix MY_tM' with any symmetric positive semi-definite function of $Y_t, Y_{t-1}, \dots, Y_{t-p+1}$. For a VAR(p) process, the equivalent of (A.3) reads:

$$E_t(Y_{t+1}) = \sum_{j=1}^p M_j Y_{t+1-j} M_j' + K\Sigma. \quad (\text{A.5})$$

In the paper, unless differently stated, we refer only to VAR(1) specifications.

B Appendix

B.1 Tables

Table 1: The different specifications of the WAR models.

full WAR(1)	M1
block diagonal WAR(1)	M2
restricted block diagonal WAR(1)	M3
diagonal WAR(1)	M4
restricted diagonal WAR(1)	M5
full HAR-WAR	M6
block diagonal HAR-WAR	M7
restricted block diagonal HAR-WAR	M8
diagonal HAR-WAR	M9
restricted diag. HAR-WAR	M10

Table 2: Grouping of the stocks according to their market sector

Stock	Tick symbol	Sector
Citigroup Inc.	C	Financial
Bank of Marica	BAC	
JP Morgan Chase Co.	JPM	
Boeing Co.	BA	Capital Goods
Caterpillar Inc.	CAT	
Fedex	FDX	
International Business Machine Corp.	IBM	Technology
Hewlett Packard Co.	HPQ	
Texas Instrument Inc.	TXN	
Kraft Food Inc.	KFT	Consumer Service
Procter Gamble Co.	PG	
Time Warner Inc.	TWX	

Table 3: Amisano-Giacomini test. Gray cells indicate a 1% preference for the column model or for the row model, according to the sign, negative or positive

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
M1		-3.62	-4.16	-0.35	0.28	0.84	-2.58	-3.03	-0.16	0.46
M2			-1.85	1.97	3.83	2.87	0.47	-0.02	2.33	3.44
M3				2.22	4.33	3.08	0.93	0.56	2.55	3.76
M4					0.53	0.92	-1.61	-1.85	0.30	0.66
M5						0.57	-2.58	-3.03	-0.32	0.38
M6							-2.85	-3.05	-0.92	-0.44
M7								-1.57	2.35	2.62
M8									2.59	3.00
M9										0.47
M10										

Table 4: Amisano-Giacomini test p -values. A yellow cell indicates a 1% preference for the column model, a red cell indicates a 1% preference for the row model (test is row model minus column model).

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
M1		0.00	0.00	0.36	0.39	0.20	0.00	0.00	0.44	0.32
M2			0.03	0.02	0.00	0.00	0.32	0.49	0.01	0.00
M3				0.01	0.00	0.00	0.18	0.29	0.01	0.00
M4					0.30	0.18	0.05	0.03	0.38	0.25
M5						0.28	0.01	0.00	0.37	0.35
M6							0.00	0.00	0.18	0.33
M7								0.06	0.01	0.00
M8									0.00	0.00
M9										0.32
M10										

Table 5: Values of the Diebold-Mariano test for the comparison of the loss-differentials of loss function 1

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
M1		1.41	1.37	1.52	1.22	1.46	1.48	1.45	1.48	1.34
M2			1.17	-1.20	0.33	-0.84	0.64	1.27	-0.74	1.00
M3				-1.19	-0.88	-0.93	-0.21	0.70	-0.88	0.33
M4					0.90	-0.06	1.44	1.39	0.68	1.18
M5						-0.70	0.15	0.79	-0.60	0.69
M6							1.20	1.19	1.08	0.98
M7								1.15	-1.23	0.42
M8									-1.21	-0.91
M9										0.94
M10										

Table 6: P-values of the Diebold-Mariano test for the comparison of the loss-differentials of loss function 1

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
M1		0.08	0.09	0.06	0.11	0.07	0.07	0.07	0.07	0.09
M2			0.12	0.12	0.37	0.2	0.26	0.1	0.23	0.16
M3				0.12	0.19	0.18	0.42	0.24	0.19	0.37
M4					0.18	0.47	0.08	0.08	0.25	0.12
M5						0.24	0.44	0.22	0.27	0.25
M6							0.12	0.12	0.14	0.16
M7								0.13	0.11	0.34
M8									0.11	0.18
M9										0.17
M10										

Table 7: Values of the Diebold.Mariano test for the comparison of the loss-differentials of loss function 2. A grey cell indicates preference for the column model (positive sign) or for the row model (negative sign).

M1	7.22	6.99	3.37	-1.40	-0.13	6.41	6.44	3.85	-1.70
M2		1.63	-5.15	-4.31	-4.47	2.49	2.69	-0.30	-4.53
M3			-4.93	-4.24	-4.62	2.40	2.67	-0.46	-4.44
M4				-2.89	-1.23	4.43	4.48	2.26	-3.16
M5					0.72	4.18	4.20	3.18	-2.36
M6						5.43	5.50	3.46	-0.9
M7							2.24	-1.51	-4.39
M8								-1.62	-4.41
M9									-3.38
M10									

Table 8: P-values of the Diebold.Mariano test for the comparison of the loss-differentials of loss function 2. A red cell indicates preference for the row, a yellow for the column.

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
M1		0.00	0.00	0.00	0.08	0.45	0.00	0.00	0.00	0.04
M2			0.05	0.00	0.00	0.00	0.01	0.00	0.38	0.00
M3				0.00	0.00	0.00	0.01	0.00	0.32	0.00
M4					0.00	0.11	0.00	0.00	0.01	0.00
M5						0.24	0.00	0.00	0.00	0.01
M6							0.00	0.00	0.00	0.18
M7								0.01	0.07	0.00
M8									0.05	0.00
M9										0.00
M10										

Table 9: Model Confidence Set P-Values for Loss Function 1 and 2

Loss Function 1				Loss Function 2			
MCS test: R		MCS test: SQ		MCS test: R		MCS test: SQ	
Model	P-value	Model	P-value	Model	P-value	Model	P-value
1	0.276	1	0.200	1	0.000	1	0.000
4	0.290	4	0.351	10	0.001	10	0.001
6	0.417	6	0.365	6	0.001	6	0.001
2	0.417	2	0.365	4	0.001	4	0.005
9	0.417	9	0.365	5	0.011	5	0.008
5	0.417	5	0.554	2	0.063	2	0.031
7	0.417	7	0.554	9	0.063	9	0.031
10	0.546	10	0.571	3	0.063	3	0.031
3	0.546	3	0.571	7	0.063	7	0.031
8	1.000	8	1.000	8	1.000	8	1.000