KIER DISCUSSION PAPER SERIES

KYOTO INSTITUTE
OF
ECONOMIC RESEARCH

Discussion Paper No.778

"Ranking multivariate GARCH models
by problem dimension: An empirical evaluation"

Michael McAleer and Massimiliano Caporin

June,2011

KYOTO UNIVERSITY

KYOTO, JAPAN

Ranking Multivariate GARCH Models by Problem Dimension: An Empirical Evaluation*

Massimiliano Caporin

Department of Economics and Management "Marco Fanno"
University of Padova
Italy

Michael McAleer

Econometric Institute
Erasmus School of Economics
Erasmus University Rotterdam
The Netherlands
and
Tinbergen Institute
The Netherlands
and
Institute of Economic Research
Kyoto University
and
Department of Quantitative Economics
Complutense University of Madrid

First draft: January 2009
This version: May 2011

* The authors wish to thank the Editor, Associate Editor, two referees, Christian Hafner, Sébastien Laurent, Francesco Violante, Roberto Casarin, Tommaso Proietti, Gian Piero Aielli, Adelchi Azzalini, Riccardo Jack Lucchetti, Eduardo Rossi, Giovanni Urga, participants at seminars in Louvain-la-Neuve and Zurich, and participants at the Italian Statistical Society XLV Conference, Padova, June 2010, CFE10 conference, London, December 2010, and ICEEE conference, Pisa, January 2011, for helpful comments and suggestions. For financial support, the second author wishes to thank the Australian Research Council, National Science Council, Taiwan, and the Japan Society for the Promotion of Science.

Abstract

In the last 15 years, several Multivariate GARCH (MGARCH) models have appeared in the literature. Recent research has begun to examine MGARCH specifications in terms of their out-of-sample forecasting performance. In this paper, we provide an empirical comparison of a set of models, namely BEKK, DCC, Corrected DCC (cDCC) of Aeilli (2008), CCC, Exponentially Weighted Moving Average, and covariance shrinking, using historical data of 89 US equities. Our methods follow part of the approach described in Patton and Sheppard (2009), and the paper contributes to the literature in several directions. First, we consider a wide range of models, including the recent cDCC model and covariance shrinking. Second, we use a range of tests and approaches for direct and indirect model comparison, including the Weighted Likelihood Ratio test of Amisano and Giacomini (2007). Third, we examine how the model rankings are influenced by the cross-sectional dimension of the problem.

Keywords: Covariance forecasting, model confidence set, model ranking, MGARCH, model comparison.

JEL codes: C32, C53, C52.

1. Introduction

Multivariate Volatility Models (MVM) have attracted considerable interest over the last decade. This may be associated with the increased availability of financial data, the increased computational powers of computers, and the fact that the finance industry has begun to realize the possible advantages of these models.

The recent literature on the topic has moved from the introduction of new models to the efficient estimation of existing models. Among the most highly-cited topics are the "curse of dimensionality" and "feasible model estimation". In fact, the feasibility of model estimation is now of central interest, with many studies proposing appropriate parameterizations of known models (Billio et al, 2006, Billio and Caporin, 2009, Franses and Hafner, 2009, Caporin and Paruolo, 2009, Bonato et al., 2009, and Asai et al., 2009), or focusing on special estimation methods (Engle and Kelly, 2008, Engle et al., 2008, and Fan et al., 2007).

A second strand of the literature has focused on the statistical or asymptotic properties of the models and of the proposed estimators (Comte and Liebermann, 2003, Ling and McAleer, 2003, McAleer et al. 2008, Engle et al. 2008, Aielli, 2008, Caporin and McAleer, 2011, Hafner and Preminger, 2009, and Francq and Zakoian, 2010). These studies have noted that only in special cases are the asymptotic properties known, and in some of them only under untestable moment restrictions, or under claimed regularity conditions (see Caporin and McAleer (2011) for a detailed discussion).

Despite the theoretical properties typically being assumed under unstated and untestable regularity conditions, many proposed models have been used widely in empirical financial studies. Within this framework, a different problem arises: How can we compare and rank models characterized by different structures? Some research has recently appeared in the literature to tackle the problem, first at the univariate level (Hansen and Lunde (2005, 2006)), then for the evaluation of alternative covariance models (Engle and Colacito (2006), Engle and Sheppard (2008), Clements et al. (2009), Patton and Sheppard (2009), and Laurent at al. (2009, 2010)). These papers have presented limited comparisons across a small range of models. Engle and Colacito (2006) compare only the DCC model of Engle (2002) against a constant correlation model, and in a datasets with a cross-sectional dimension equal to 2 (that is, two stock market or bond indices) or 34 (the same series as used in Cappiello et al., 2006). Engle and Sheppard (2008) is quite an extensive study for the model considered, but uses a single cross-sectional dimension (50 sector indices defined within the perimeter of the S&P 500 index). Patton and Sheppard (2009) is a theoretical contribution on the approaches to be used for the evaluation of covariance forecasts, and does not include an empirical application (even with low cross-sectional dimensions) showing the arguments for and against the various methods. Clements et al. (2009) focus on dynamic correlation models, and present results for a cross-sectional dimension equal to 5 (five US based future contracts). Laurent et al. (2009) focus on the consistency of multivariate loss functions, report an empirical example over three assets, and simulations for a bivariate case. Laurent et al. (2010) consider a moderately large set of models, but focus on a 10 assets example, and place emphasis on the model accuracy against a DCC benchmark. Furthermore, all of the previous papers include the DCC model of Engle (2002), and are thereby exposed to the estimation (in)consistency problems discussed in Aielli (2008).

The methods of comparison used in the previous contributions could be viewed as two large classes (see Patton and Sheppard (2009)), namely the direct and indirect evaluation of volatility forecasts. The first group includes the Mincer-Zarnowitz regression (Mincer and Zarnowitz, 1969), Diebold-Mariano test (Diebold and Mariano, 1996, and West, 1996, 2006), Reality Check of White (2000), Superior Predictive Ability (SPA) test of Hansen (2005), and the Model Confidence Set (MCS) approach of Hansen et al. (2003, 2010). The second group includes approaches based on the comparison of loss functions adapted to the needs of covariance forecasts. This is the case, for instance, of asset allocation and risk management, where loss functions could be defined using global minimum variance portfolios returns, such as in Engle and Colacito (2006) and Patton and Sheppard (2009), or within a Value-at-Risk framework, as in Ferreira and Lopez (2005).

The tests that compare directly the covariance forecasts fit the general framework of loss-function comparison, as discussed in Clements et al. (2009) and Patton and Sheppard (2009). The Diebold-Mariano and West approaches are valid for pairwise comparisons of the models, the Reality check and SPA require the identification of a benchmark model, whereas MCS does not require a benchmark specification. Overall, the MCS approach seems to be preferred, and is the most appropriate as it provides a statistical test and a method for determining which models are statistically equivalent with respect to a given loss function. Despite the use of a bootstrap method for the evaluation of test statistic, MCS is computationally feasible, efficient and statistically robust. With respect to the indirect comparison of volatility forecasts, an interesting result has been shown in Clements et al. (2009), that illustrates how utility-based loss functions (in particular, quadratic utilities) make the impact of the covariance model very modest. The approach of Engle and Colacito (2006) should provide interesting results, even for large cross-sectional dimensions.

Working in a purely empirical setting, in this paper we contribute to the literature on covariance forecast evaluation in several ways. First, our selection of models to be compared differs from those of previous studies. Similar to the literature, we consider the CCC model of Bollerslev (1990), DCC model of Engle (2002), Scalar BEKK model with targeting of Ding and Engle (2002), the OGARCH model of Alexander (2001a,b), and the naïve Exponentially Weighted Moving Average approach. We complement this set by including the cDCC model of Aielli (2008), and the covariance shrinking approach of Ledoit and Wolf (2003, 2004).

The introduction of the cDCC model allows evaluation of the impact of both the lack of consistency and the existence of bias in the estimated parameters of the DCC model of Engle (2002). Aielli (2008) shows that the bias depends on the persistence of the DCC dynamic parameters. We are interested in evaluating if DCC could be used, regardless of its inconsistency. This fact is of interest as DCC has been proposed as a model with correlation targeting, whereas cDCC cannot be targeted, as discussed in Caporin and McAleer (2011). By including the covariance shrinking method, we evaluate its advantages in large cross-sectional dimensions. Covariance shrinking is computationally feasible and may also reduce the problems associated with the inversion of large covariance matrices, wherein inversion could be unstable due to the presence of small eigenvalues in the empirical covariances. Furthermore, the presence in the model set of the Scalar BEKK allows determining if the separated estimation of variances and correlations (typical of CCC- and DCC-type models) is to be preferred to the joint estimation of the entire covariance (as in BEKK-type models). Such an analysis could provide a confirmation of the result of Zumbach (2009) that shows evidence of a preference for covariance models with respect to variance and correlation specifications.

Second, we use the weighted likelihood ratio test of Amisano and Giacomini (2007), which uses loss-function comparisons of equal predictive ability based on a likelihood loss function. The test will be applied both in the direct evaluation of covariance forecasts and as an alternative to the Diebold-Mariano test. An advantage of the Amisano and Giacomini (2007) approach is that the test statistic is not a function of the true and unknown covariance matrix. As a result, the test is not affected by the estimation error implicit in the use of covariance proxies. The latter element will be further investigated on different loss functions by contrasting the results with a nosiy proxy to those with a realized covariance proxy, with the purpose of extending the results of Hansen and Lunde (2005, 2006), and completing those in Laurent at al. (2009, 2010).

Third, we will evaluate and rank the alternative models over different cross-sectional dimensions, starting from five assets, and up to 89 assets, which we select from the S&P100 constituents.[1] We will determine if the cross-sectional dimension has a role in determining the preference ordering across models. In other words, by comparing models over an increasing number of variables, we will examine if estimation error and model error play a role in the forecasts of conditional covariance models. The financial literature has discussed extensively the impact of estimation error for the mean returns, leading to results suggesting its strong impact, and making naïve allocations preferable to optimal allocations because of the reduced impact of estimation error (see De Miguel et al., 2009). We draw a parallel within the MGARCH model set, and attempt to answer the question: If the cross-sectional dimension is large, does estimation error affect model performance? If simple or naïve models are preferred, we could interpret this as preliminary evidence in this direction.

We stress that we are comparing alternative feasible models for the evaluation of conditional covariance and/or correlation matrices. The models we consider all belong to the GARCH and Dynamic Conditional Correlation families, thereby excluding Multivariate Stochastic Volatility models. From our perspective, these models, despite being theoretically appealing, suffer for the curse of dimensionality in a stronger way

[1] A similar dataset has been used in Engle et al. (2008).

than MGARCH specifications, and their estimation in large cross-sectional dimensions is likely to be even more complicated than the models considered in this paper. For surveys of Multivariate Stochastic Volatility models, see Asai et al. (2006), and Chib et al. (2009).

Furthermore, we focus on extremely simple models (all are scalar representations), and follow the quasi-maximum likelihood estimation approach. We do not consider more complex parameterizations because the emphasis is on simplicity. We are not interested in the determination of an optimal model or estimation method, but rather on baseline specifications, namely those that are the most common among practitioners, and try to verify if they are equivalent. Clearly, 89 assets is far from the traditional problem dimension of large portfolio managers, but this sheds some light on a comparison of model performance across an increasing number of assets. Finally, we stress that our focus in on empirical application, and is not intended to provide a methodological contribution to the most appropriate methods for model comparison, which will be left for future research based on a simulation approach.

Our results show that the use of a realized covariance proxy has a relevant impact on model rankings. Furthermore, the rankings are not greatly affected by the problem size, and they stabilize as the number of assets starts to increase. Across the models, some preference may go to the DCC-type and OGARCH-type specifications, while the naïve specifications are generally found to be underperforming. Finally, given the previous comment, we do not find a confirmation of Zumbach (2009) for a preference of covariance models compared to variance and correlation models.

The paper proceeds as follow. Section 2 presents the model, briefly discusses the issue of covariance and correlation targeting, and shows the specifications to be estimated. Section 3 discusses the methods and approaches used to compare the models. Section 4 presents the dataset used and reports the empirical results. Section 5 gives some concluding comments.

2. Feasible covariance and correlation models for large cross-sectional dimensions

This section briefly introduces the models that will be compared in the empirical application. Let x_t denote a k-dimensional vector of financial variables (returns), μ_t represent the expected mean of x_t from a conditional mean model, and ε_t the mean innovation vector. The following relations hold:

$$x_t - \mu_t = \varepsilon_t \mid I^{t-1} \sim D\left(0, \Sigma_t\right) \tag{1}$$

where I^{t-1} is the information set at time *t-1*, *D(.)* denotes a multivariate density, and Σ_t represents the covariance matrix that is determined conditionally on the information set at time *t-1*.

In the following, we do not consider the effects of different mean specifications. The mean is fixed at the sample mean determined over the same sample used for the estimation of the parameters, such that $\hat{\mu}_t = \frac{1}{t-1}\sum_{i=1}^{t-1} x_i$. The mean could be based on a variety of time series or financial models, which are not the main concern of this paper. What is relevant is that, for each pair of covariance models that is compared, the mean models are identical. As a result, all forecast discrepancies are due to differences in the expected covariances, while all in-sample differences are due to differences in the estimated covariance models.

2.1 Scalar BEKK

The first model we estimate is the Scalar BEKK with targeting constraint (see Engle and Kroner, 1995; Ding and Engle, 2001; Caporin and McAleer, 2008, 2011)[2]. The model is given as

[2] Note that the VECH model adopted by Engle and Sheppard (2008) is equivalent to a scalar BEKK model. We do not consider the VECH model class in the following as the feasible VECH generally has a corresponding BEKK representation.

$$\Sigma_t = \bar{\Sigma} + \alpha\left(\varepsilon_{t-1}\varepsilon_{t-1}' - \bar{\Sigma}\right) + \beta\left(\Sigma_{t-1} - \bar{\Sigma}\right) \tag{2}$$

where α and β are scalar coefficients and the matrix $\bar{\Sigma} = E\left[\varepsilon_t\varepsilon_t'\right]$ is estimated using the sample covariance. Scalar BEKK in (2) is feasible[3] even for very large cross-sectional dimensions as it contains only two parameters that must be estimated by maximum likelihood, namely the parameters driving the model dynamics. Notably, the Scalar BEKK model has standard asymptotic properties, as shown by, for example, Hafner and Preminger (2009) under the existence of 6th order moments of the process[4].

We also consider a generalization of the scalar BEKK which includes asymmetry, the different impact of shocks on conditional variances and covariances depending on the shock sign. The Scalar BEKK with asymmetry (ABEKK) is given as

$$\Sigma_t = \bar{\Sigma} + \alpha\left(\varepsilon_{t-1}\varepsilon_{t-1}' - \bar{\Sigma}\right) + \beta\left(\Sigma_{t-1} - \bar{\Sigma}\right) + \delta\left(\eta_{t-1}\eta_{t-1}' - \bar{\mathrm{N}}\right) \tag{3}$$

where $\eta_{i,t} = \varepsilon_{i,t} I\left(\varepsilon_{i,t} < 0\right)$, $I(\cdot)$ is the indicator function, and $\bar{\mathrm{N}} = E\left[\eta_t\eta_t'\right]$.[5]

2.2 Variance and correlation models

We will estimate three models based on a decomposition of the covariance matrices into variances and correlations. The first is the CCC model of Bollerslev (1990)[6] which, starting from (1), assumes that the covariance matrix satisfies

$$\Sigma_t = D_t R D_t \tag{4}$$

where $D_t = diag\left(\sigma_{1,t}, \sigma_{2,t}, \ldots \sigma_{k,t}\right)$ is a diagonal matrix of conditional standard deviations, and R is an unconditional correlation matrix[7]. We first assume that all the conditional variances follow a simple GARCH(1,1) process without asymmetry in order to make the model directly comparable with Scalar BEKK. Second, we consider a CCC model where the variances follow the GJR-GARCH(1,1) model of Glosten et al. (1993),

[3] The diagonal BEKK and VECH parameterizations are not considered as they are not feasible for large cross-sectional dimensions.
[4] See also Jeantheau (1998), Comte and Lieberman (2003), and McAleer et al. (2009)
[5] This specification is identical to the Asymmetric VECH model adopted by Engle and Sheppard (2008).
[6] The CCC model is a special case of the VARMA-GARCH model of Ling and McAleer (2003).
[7] The operator *diag(a)* generates a diagonal matrix, with the vector *a* along the main diagonal.

thereby including asymmetry in the variances (we call this model CCC-GJR). As distinct from Laurent et al. (2010), we do not consider a wider set of univariate models, in order to avoid overfitting (it is difficult to have long memory over the entire set of series, or to have models with orders greater than 1 over all the assets).

The model is estimated using a two-step approach, namely the conditional variances on each specific series at first, and then the unconditional correlation matrix is estimated using the sample estimator over the standardized residuals $\hat{D}_t^{-1}\varepsilon_t$ [8]. This approach makes the model feasible, even for a large number of assets.

Engle (2002) and Tse and Tsui (2002) proposed two generalizations of the CCC model, where the constant correlation matrix in (4) is replaced by a time-varying conditional correlation. We consider here the DCC model of Engle (2002), which is given as:

$$\Sigma_t = D_t R_t D_t \tag{5}$$

$$R_t = \bar{Q}_t^{-1} Q_t \bar{Q}_t^{-1}, \qquad \bar{Q}_t = diag\left(dg\left(Q_t\right)\right)^{½} \tag{6}$$

$$Q_t = S + \alpha\left(D_{t-1}^{-1}\varepsilon_{t-1}\varepsilon_{t-1}{}' D_{t-1}^{-1} - S\right) + \beta\left(Q_{t-1} - S\right) \tag{7}$$

where D_t is the same as for the CCC model, S is a parameter matrix, and α and *ß* are the scalar parameters driving the model dynamics[9]. Following Engle (2002), the model is estimated with a three-stage approach, namely estimate the conditional variance parameters and filter them, estimate S as the correlation matrix of the standardized residuals $\hat{D}_t^{-1}\varepsilon_t$ by means of a sample estimator, and then, conditionally on the previous estimates, maximize the conditional correlation log-likelihood with respect to the parameters driving the dynamics in (7). The introduction of a multi-step estimation method clearly reduces the efficiency, as shown in Engle and Sheppard (2001), but makes the model feasible with large cross-sectional dimensions.

The model in (5)-(7) theoretically includes targeting, as defined in Caporin and McAleer (2011), but only under assumptions which are analysed and criticized in Aielli (2008). Without targeting, the model is inextricably exposed to the curse of dimensionality, as the matrix S contains $0.5k(k-1)$ parameters to be jointly estimated with α and *ß*. Similarly to the CCC, we consider two possible cases for the DCC model,

---

[8] The hat denotes estimated quantities.
[9] The operator *dg(A)* extracts the main diagonal from matrix *A*.

the first with GARCH(1,1) variances (DCC), while for the second we use the GJR-GARCH (DCC-GJR). Furthermore, we allow for asymmetry in the correlation process, which is given as

$$Q_t = S + \alpha\left(D_{t-1}^{-1}\varepsilon_{t-1}\varepsilon_{t-1}{}' D_{t-1}^{-1} - S\right) + \beta\left(Q_{t-1} - S\right) + \delta\left(\upsilon_{t-1}\upsilon_{t-1}{}' - \bar{\Upsilon}\right) \quad (8)$$

where $\upsilon_{i,t} = \varepsilon_{i,t}\sigma_{i,t}^{-1} I\left(\varepsilon_{i,t}\sigma_{i,t}^{-1} < 0\right)$, and $\bar{\Upsilon} = E\left[\upsilon_t\upsilon_t{}'\right]$. The resulting model will be referred to as ADCC if the conditional variances follow a GARCH(1,1) process, or ADCC-GJR if we also include asymmetry in the conditional variances.

Aielli (2008) shows that the sample estimator of *S* used in the second step of the DCC estimation method is inconsistent, thereby also affecting the consistency of the third step. In order to resolve this serious issue, Aielli (2008) introduces the cDCC model, which replaces (7) with

$$\begin{aligned} Q_t &= S + \alpha\left(\tilde{Q}_{t-1}^{1/2} D_{t-1}^{-1}\varepsilon_{t-1}\varepsilon_{t-1}{}' D_{t-1}^{-1}\tilde{Q}_{t-1}^{1/2} - S\right) + \beta\left(Q_{t-1} - S\right) \\ \tilde{Q}_{t-1} &= diag\left(dg\left(Q_{t-1}\right)\right) \end{aligned} \quad (9)$$

where the parameter matrix *S* is symmetric, has unit elements over the main diagonal, and is now the covariance matrix of the innovations $\tilde{Q}_t^{1/2} D_t^{-1}\varepsilon_t$, which are not observable. The modification restores consistency, under unstated assumptions, but again exposes the model to the curse of dimensionality as the matrix *S* in (8) has to be estimated (see Aielli (2008) for further details). As noted in Caporin and McAleer (2011), correlation targeting is excluded for the cDCC model as *S* is not a correlation matrix, and is not estimated using the available sample information. Aielli (2008) suggests a feasible estimation method that is similar to the profile likelihood.

We note that Aielli (2008) shows that the lack of consistency of the three-step DCC estimator depends strictly on the persistence of the parameters driving the correlation dynamics and on the relevance of the innovations. The bias is an increasing function of both α and $\alpha + \beta$. Not surprisingly, the typical parameter estimates obtained from fitting DCC models are small, and are close to 0 for α and to 1 for $\alpha + \beta$, thereby leading to an opposite effect on the size of the bias. Therefore, in this paper we will determine if the bias is relevant in practical applications as a commentary on the inconsistent estimates of the standard scalar DCC model. With a notation similar to that adopted for the DCC, we label Aielli's (2006) model as cDCC if the variances follow a GARCH(1,1) process and cDCC-GJR if the conditional variances include asymmetry.[10]

2.3 Factor GARCH

Factor GARCH is a model class including two subgroups: in the first set we have specifications where the factors are latent, such as Engle et al. (1990) and Lanne and Saikkonen (2007); the second group includes models where the multivariate structure arises from linear combinations of univariate GARCH models, such as in Alexander (2001a,b), Vrontos et al. (2003), and Van der Weide (2002). Further details on this model class can be found in Bauwens et al. (2006). As a competitor to BEKK and dynamic conditional correlation models, we consider here the OGARCH model of Alexander (2001a,b). We motivate the choice by the simplicity of the model compared with the alternative specifications mentioned above, which could also be influenced by the curse of dimensionality.

In the OGARCH model the covariance matrix is represented as:

$$\Sigma_t = DP\Lambda^{1/2} H_t \Lambda^{1/2} P' D \qquad (10)$$

where D is the diagonal matrix of unconditional standard deviations of ε_t, P is the matrix of eigenvectors, and Λ is the diagonal matrix of eigenvalues obtained from the unconditional correlation matrix of $D^{-1}\varepsilon_t$, and H_t is the diagonal matrix of the principal components conditional variances. The principal components are given as $p_t = P^{-1}\Lambda^{-1/2}\left(D^{-1}\varepsilon_t\right)$, and they could all follow either GARCH(1,1) processes (in this case, the model is denoted as OGARCH) or GJR processes (we label this model OGARCH-GJR).

2.4 Naïve specifications

[10] For the cDCC model we do not consider an extension including asymmetry in the correlations as the model cannot be estimated using the approach proposed by Aielli (2008). As the main purpose of this paper is not the development of new models and estimation methods, we leave the introduction of the Asymmetric cDCC to future research.

The last two models considered are the Exponentially Weighted Moving Average model and the Covariance Shrinking approach of Ledoit and Wold (2003, 2004). The EWMA model provides a recursion for the evaluation of the conditional covariance matrix, which is based on a single parameter λ:

$$\Sigma_t = (1-\lambda)\varepsilon_{t-1}\varepsilon_{t-1}' + \lambda\Sigma_{t-1} \quad (11)$$

In the empirical application, contrary to standard practice, we estimate the parameter λ, called the smoothing coefficient, as it requires limited computational effort. By construction, the EWMA is feasible even for very large cross-sectional dimensions.

Finally, we consider the covariance shrinking approach of Ledoit and Wolf (2003, 2004). The authors proposed a method that is designed to find a compromise between the large estimation errors in the sample covariance and the misspecification error in the estimators of the covariance. They suggest determining a covariance by combining a sample estimator of the covariance and a single index covariance (Ledoit and Wolf, 2003), or a constant correlation covariance (Ledoit and Wolf, 2004). Following the covariance shrinking approach, we define the expected covariance for time *t* as follows:

$$\Sigma_t = (1-\lambda)S_{t-1} + \lambda F_{t-1} \quad (12)$$

where S_{t-1} is the sample covariance matrix determined up to time *t-1*, and F_{t-1} is a structured estimator determined using the information set to time *t-1*, and is called the shrinkage target. The coefficient, λ, which is the shrinkage constant, has to be estimated, and depends on the form of the shrinkage target (for further details, see Ledoit and Wolf (2003, 2004)). In the following, we will consider as the shrinkage target the covariance with constant correlation, as described in Ledoit and Wolf (2004).

3. Comparing competing covariance and correlation models

We will present briefly the approaches to be used in comparing the models described in the previous section. Before moving to the methods, we introduce some notation.

It is assumed that the models are to be compared using out-of-sample forecasts, where forecasts are made one period ahead and for an evaluation period from *T+1* to *T+h*. Information to time *T* is used to estimate the various models and to produce the conditional forecasts for time *T+1*. The estimation sample is rolled forward, and information from time *2* to *T+1* is used to forecast the covariance matrix for time *T+2*, and so on, to time *T+h*. In order to avoid any dependence on the mean dynamics, we fit the mean using its sample estimator across all models (the sample mean is estimated with the same rolling approach). The one-step-ahead covariance forecasts for time *T+i* are denoted by $\hat{\Sigma}_{T+i}^{m}$, where *m* is the model index (*m=1,2,…M*). Note that, by construction, the forecasts are conditional on the information set at time *T+i-1*. The mean forecasts are denoted by $\hat{\mu}_{T+i}$, and do not depend on the model. For simplicity, we suppress the conditioning information set from the forecast notation.

We follow Patton and Sheppard (2009) and consider separately the direct and indirect evaluation methods.

3.1 Direct model evaluation methods

Within the first group, we include approaches based on the use of loss functions, namely the Diebold-Mariano test, the test proposed by Amisano and Giacomini (2007), and the MCS approach of Hansen et al. (2003 and 2010).

Let us denote a loss function for time *T+i* and model *l* as $lf_{l,T+i}$. Then, the test for equal predictive ability between two competing models corresponds to checking the null hypothesis of zero loss function differentials, $H_0 : E\left[\overline{lf}_j - \overline{lf}_l\right] = E\left[\overline{LF}_{jl}\right] = 0$, where *l* and *j* are two different model indices, $\overline{lf}_j = h^{-1}\sum_{i=1}^{h} lf_{j,T+i}$, and $\overline{LF}_{jl} = \overline{lf}_j - \overline{lf}_l$. In this setting, the test statistic is given as

$$t_{jl} = \frac{\sqrt{h}\,\overline{LF}_{jl}}{Var\left(\sqrt{h}\,\overline{LF}_{jl}\right)^{½}} \sim N(0,1) \qquad (13)$$

where $Var\left(\sqrt{h}\,\overline{LF}_{jl}\right)$ is a heteroskedasticity and autocorrelation (HAC) consistent estimate of the asymptotic variace of $\sqrt{h}\,\overline{LF}_{jl}$. If the null hypothesis of equal forecasting ability is rejected, the test statistic sign suggests model preference: positive (negative) values indicates a preference for the second (first) model as it provides smaller losses.

We consider the two loss functions reported below:

i) $lf^{a}_{m,T+i} = \frac{1}{k^2}\left(vec\left(\hat{\Sigma}^{m}_{T+i} - \tilde{\Sigma}_{T+i}\right)' vec\left(\hat{\Sigma}^{m}_{T+i} - \tilde{\Sigma}_{T+i}\right)\right),$ (14)

ii) $lf^{b}_{m,T+i} = \log\left|\hat{\Sigma}^{m}_{T+i}\right| + e_{T+i}{}' \left(\hat{\Sigma}^{m}_{T+i}\right)^{-1} e_{T+i},$ (15)

where $e_{T+i} = x_{T+i} - \hat{\mu}_{T+i}$ (note that the observed time $T+i$ return is used), and the time $T+i$ true volatility is approximated by a proxy, $\tilde{\Sigma}_{T+i}$. In the empirical evaluation, we will use two different choices of the volatility proxy, $\tilde{\Sigma}_{T+i}$: a first possibility is given by the cross-product of mean forecast errors e_{T+i}, so that $\tilde{\Sigma}_{T+i} = e_{T+i}e_{T+i}{}'$. However, this is a noisy proxy, as shown in Patton and Sheppard (2009) and Laurent et al. (2010), among others. As an alternative, we consider a realized covariance estimator. Within the class of possible approaches, we choose the Multivariate Realized Kernel of Barndorff-Nielsen et al. (2008), with data sincronyzed at the 5-minute frequency.[11]

The first function, equation (14), corresponds to the Mean Squared Error (MSE) loss adopted in the Diebold-Mariano test. The MSE loss function belongs to the class of loss functions defined in Patton and Sheppard (2009)[12] that are robust to the noise in the volatility proxy used.

The second loss function corresponds to minus the logarithmic scores, and makes the test statistic equivalent to the Amisano and Giacomini (2007) weighted likelihood ratio test when all points over the forecast horizon have identical weight. We stress that this loss function does not depend on a volatility proxy, and so is not exposed to the estimation error of the underlying and unknown true volatility. Furthermore, it evaluates

[11] For details on the estimator adopted, the interested reader should refer to Barndorff-Nielsen et al. (2008). Alternative approaches are discussed in Andersen et al. (2003, 2009) and Barndorff-Nielsen and Shephard (2004), among others.

[12] See also Clements et al. (2009) and Laurent et al. (2009). Patton and Sheppard (2009) also consider the QLIKE loss function of Patton (2010), but in the multivariate framework, the QLIKE loss function is infeasible when the volatility proxy is the cross-product of realized returns (see Laurent et al., 2009).

the fit of all models by means of a Gaussian score measured using the mean forecast errors.[13]

The previous tests permit a pairwise comparison of models. However, the test outcomes do not ensure either that an optimal model is clearly identified or that a clear model ordering is obtained. Furthemore, when dealing with multiple comparison, as in this case, a Bonferroni bound correction is needed. For these reasons, we consider the Model Confidence Set approach, which performs a joint forecast comparison across all models.

The MCS performs an iterative selection procedure, testing at step *j* the null hypothesis of equal predicting ability of all models included in a set $\mathcal{M}_\ell$ (the starting set $\mathcal{M}_0$ contains all the models) under a given loss function.

The null hypothesis has the form

$$H_0 : E\left[\overline{lf}_{jl} - \overline{lf}_{jl}\right] = E\left[\overline{LF}_{jl}\right] = 0, \quad j > l, \forall j, l \in \mathcal{M}_\ell \tag{16}$$

where the notation is the same as in (13). In order to test the null hypothesis, we use the following two test statistics proposed by Hansen et al. (2003)[14]:

$$t_R = \max_{j,l \in \mathcal{M}_\ell} \left| \frac{\overline{LF}_{jl}}{Var\left(\overline{LF}_{jl}\right)^{½}} \right| \tag{17}$$

$$t_{SQ} = \sum_{j,l \in \mathcal{M}_\ell, j>l} \left(\frac{\overline{LF}_{jl}}{Var\left(\overline{LF}_{jl}\right)^{½}} \right)^2 \tag{18}$$

where $Var\left(\overline{LF}_{jl}\right)$ is a bootstrap estimate of the variance of $\overline{LF}_{jl}$, and the p-values of the test statistics are determined using a bootstrap approach. If the null hypothesis is rejected at a given confidence level, the worst performing model is excluded from the set (rejection is determined on the basis of bootstrap p-values under the null). Such a model is identified as follows:

[13] As the mean forecasts are identical across models, the differences in the losses are solely due to differences across the covariance models.

[14] Hansen et al. (2010) include additional test statistics that are not included in this paper.

$$j = \arg\max_{j \in \mathcal{M}_\ell} \left(\sum_{l \in \mathcal{M}_\ell} \overline{LF}_{jl} \right) \left(Var\left(\sum_{l \in \mathcal{M}_\ell} \overline{LF}_{jl} \right)^{½} \right)^{-1} \tag{19}$$

where the variance is computed using a bootstrap method. In the empirical analysis given below, we will use the loss functions introduced in (14) and (15).

3.2 Indirect model evaluation methods

For the indirect evaluation of the multivariate models, we consider an asset allocation framework and compare the impact of model choice by contrasting the performances of specific portfolios: (i) equally weighted portfolio, denoted as EW, which is not exposed to the asset return mean estimation error, and is superior to many other portfolios (see De Miguel et al. (2009)); and (ii) global minimum variance portfolio, with and without short selling constraints, denoted as GMV and GMVB[15], respectively. The weights of the equally weighted portfolio are $\mathbf{w} = k^{-1}\mathbf{1}$, where $\mathbf{1}$ is a k-dimensional vector of unit elements. The GMV weights are time- and model-dependent, and are based on the covariance forecasts:

$$\mathbf{w}_{T+i}^{m} = \frac{\left(\hat{\Sigma}_{T+i}^{m}\right)^{-1}\mathbf{1}}{\mathbf{1}'\left(\hat{\Sigma}_{T+i}^{m}\right)^{-1}\mathbf{1}} \tag{20}$$

Finally, GMVB weights $\widehat{\mathbf{w}}_{T+i}^{m}$ are determined by solving the optimum problem:

$$\begin{aligned} &\arg\min_{\mathbf{w}} \ \mathbf{w}'\hat{\Sigma}_{T+i}^{m}\mathbf{w} \\ &\text{s.t. } w_l \geq 0, \quad l = 1, 2, \ldots k \\ &\text{and } \mathbf{w}'\mathbf{1} = 1 \end{aligned} \tag{21}$$

for each forecast evaluation period and for each model.

We then define the following quantities for the three portfolio strategies (based on the weights EW, GMV and GMVB, respectively)[16]:

---

[15] Where B denotes Bounded.

[16] Note that the portfolio weights for GMV and GMVB are always estimated on the basis of given model covariance forecasts. The weight vector does not report in those cases a hat to simplify the notation. The weigths of th EW strategy are not a function of the model and, as a consequence, the realized portfolio

(a) realized portfolio returns:

$$R^m_{T+i,EW} = \mathbf{w}' x_{T+i},\ R^m_{T+i,GMV} = \mathbf{w}^{m\ \prime}_{T+i} x_{T+i},\ R^m_{T+i,GMVB} = \widehat{\mathbf{w}}^{m\ \prime}_{T+i} x_{T+i},\ i = 1,2,3\ldots h,\ m = 1,2,3,\ldots M\ ;$$

(b) expected portfolio returns[17]:

$$\hat{R}^m_{T+i,EW} = \mathbf{w}' \hat{\mu}_{T+i},\ \hat{R}^m_{T+i,GMV} = \mathbf{w}^{m\ \prime}_{T+i} \hat{\mu}_{T+i},\ \hat{R}^m_{T+i,GMVB} = \widehat{\mathbf{w}}^{m\ \prime}_{T+i} \hat{\mu}_{T+i},\ i = 1,2,3\ldots h,\ m = 1,2,3,\ldots M\ ;$$

(c) realized portfolio variances[18]:

$$s^m_{T+i,EW} = \mathbf{w}' \tilde{\Sigma}_{T+i} \mathbf{w},\ s^m_{T+i,GMV} = \mathbf{w}^{m\ \prime}_{T+i} \tilde{\Sigma}_{T+i} \mathbf{w}^m_{T+i}\ ,\ s^m_{T+i,GMVB} = \widehat{\mathbf{w}}^{m\ \prime}_{T+i} \tilde{\Sigma}_{T+i} \widehat{\mathbf{w}}^m_{T+i}\ ,\ i = 1,2,3\ldots h,$$
$$m = 1,2,3,\ldots M\ ;$$

(d) expected portfolio variances:

$$\hat{s}^m_{T+i,EW} = \mathbf{w}' \hat{\Sigma}^m_{T+i} \mathbf{w},\ \hat{s}^m_{T+i,GMV} = \mathbf{w}^{m\ \prime}_{T+i} \hat{\Sigma}^m_{T+i} \mathbf{w}^m_{T+i},\ \hat{s}^m_{T+i,GMVB} = \widehat{\mathbf{w}}^{m\ \prime}_{T+i} \hat{\Sigma}^m_{T+i} \widehat{\mathbf{w}}^m_{T+i},\ i = 1,2,3\ldots h,$$
$$m = 1,2,3,\ldots M\ .$$

We note that the weights used in (a)-(d) for the GMV and GMVB strategies are always those estimated on the basis of the covariance forecasts. We do not follow Voev (2009) which considered in the realized returns and variances the optimal weights obtained by the true and unknown covariance (in our case, it would have been replaced by a proxy). In fact, that approach mixes the estimation error of the covariances with that of the portfolio weights, thereby adding a further source of uncertainty.[19] In the quantities we consider, the only difference between the expected and realized quantities is given by the covariance matrix.

---

returns and the realized portfolio variances under the EW strategy will be independent of the model used to forecast the conditional covariances.

[17] A hat is used to identify expected quantities.

[18] We use “s” and not the Greek sigma squared to denote portfolio variances to avoid possible confusion with the asset variances.

[19] Portfolio weights estimation error is, in reality, a bi-product of the covariance forecast error. This, in turn, is a function of parameter estimation, specification and model errors. Furthermore, GMVB weights might also suffer from optimization errors.

Using the quantities in (a)-(d) above, we test the null hypothesis of equal predictive ability across pairs of models at the portfolio level by using the test statistic defined in (13) and the following loss functions[20]:

i) $lf^{1}_{m,EW,T+i}=\frac{1}{2}\ln\left(\hat{s}^{m}_{T+i,EW}\right)+\frac{1}{2}\left(R_{T+i,EW}-\hat{R}_{T+i,EW}\right)^{2}\left(\hat{s}^{m}_{T+i,EW}\right)^{-1}$, (22)

ii) $lf^{2}_{m,p,T+i}=\left(\hat{s}^{m}_{T+i,p}-s^{m}_{T+i,p}\right)^{2}$, (23)

iii) $lf^{3}_{m,p,T+i}=\log\left(\hat{s}^{m}_{T+i,p}\right)+s^{m}_{T+i,p}\left(\hat{s}^{m}_{T+i,p}\right)^{-1}$, (24)

where *p=EW,GMV,GMVB.*

The loss functions in (22)-(24) are also used for the joint forecast comparison by means of the MCS approach. The loss functions in (23) and (24) are the univariate MSE and QLIKE loss functions (see Patton and Sheppard (2009)). Differently, equation (22) represents minus the logarithmic score when the mean and variances forecasts are made conditionally on time *T+i-1*. Note that the logarithmic score is evaluated at the true observed values at time *T+i*. Such a quantity can be evaluated only for the EW portfolio strategy, which is the only strategy that provides a 'true' value $R_{T+i,EW}$ [21].

Finally, in the indirect comparison, we also consider some of the model comparison approaches suggested in Engle and Colacito (2006). In particular, we report the out-of-sample averages of the expected variances $\bar{\hat{s}}^{m}_{p}=h^{-1}\sum_{i=1}^{h}\hat{s}^{m}_{T+i,p}$ (with a preference for a lower average variance), and we also test the significance of the intercept of the regressions in:

$$\frac{s^{m}_{T+i,p}}{\hat{s}^{m}_{T+i,p}}-1=\beta_{m,p}+\xi_{T+i} \tag{25}$$

[20] Note that we do not compare multivariate models indirectly by mean of utility-based loss functions because Clements et al. (2009) show that these functions make the impact of the models very limited, thereby reducing the possibility of detecting discrepancies across models.

[21] In fact, the GMV and GMVB strategies allow a determination of the realized returns, but these are exposed to the estimation error implicit in the determination of portfolio weights. As a result, in order to avoid introducing distortions in the test statistics, we consider only the EW strategy.

where ξ_{T+i} is an innovation term, and robust HAC standard errors are required. If we consider (25), accurate models should have a zero intercept[22]. Note that these comparisons are made with respect to different covariance models under the same portfolio allocation strategy. As a result, we compare alternative models (different choices of *m*) by mean of $\bar{\hat{s}}_p^m$ and $\hat{\beta}_{m,p}$ under the same portfolio strategy *p*.

4. Data description and selected models

In order to compare the models presented in the previous sections, we have selected a dataset similar to that of Engle et al. (2009). We downloaded from Datastream the S&P100 constituents at the end of March 2009. Then we selected only those assets with total return indices available from the beginning of 1997 to the end of March 2009. The selected period contains 3194 daily returns. The list of the 89 selected stocks is reported in Appendix A.

We fit the following models (acronyms are given in parentheses): standard Scalar BEKK with covariance targeting (BEKK) and with the addition of asymmetry (ABEKK); standard Scalar DCC with GARCH marginals and correlation targeting (DCC), with variance asymmetry (GJR-DCC), and with variance and correlation asymmetry (GJR-ADCC); Scalar cDCC with GARCH marginals and implicit correlation targeting (cDCC), and with variance asymmetry (GJR-cDCC); exponentially weighted moving average, with estimated smoothing coefficient (EWMA); constant conditional correlation model (CCC) with GARCH marginals and with GJR marginals (GJR-CCC); OGARCH with GARCH variances on the principal components (OGARCH), and with GJR on the principal components (GJR-OGARCH); finally, covariance shrinking (SHR) with constant correlation shrinkage target, as in Ledoit and Wolf (2004); giving a total of 13 models.

In estimating all of the models, we adopt a normal likelihood, thus resorting to Quasi Maximum Likelihood (QML) estimation. Despite the misspecification of the density,

[22] Engle and Colacito (2006) also propose pairwise comparisons based on Diebold-Mariano type tests. As they are closely related to the methods already described, we do not consider them in the empirical analysis below.

the use of a Gaussian density enables the multi-stage estimation approach for the CCC and DCC specifications. Using a Student t, by contrast, will not enable straightforward decomposition of the likelihood into the respective variance and correlation contributions.

We consider two different examples, namely medium scale and large scale. In the medium scale example, we consider a subset comprising 15 of the 89 assets; for those 15 assets, high frequency data are available at the 1 minute frequency (the list is included in Appendix A). In the medium scale example, the assets are ordered alphabetically, and we estimate the model for 5 to 15 assets. The medium scale empirical application allows evaluation of the impact on the covariance proxy used; we compare the model rankings obtained when the proxy is the realized covariance with those obtained when the cross product of realized asset returns is used. Furthemore, the impact of the noise on the proxy for the model rankings will also be evaluated over the different allocation strategies in order to examine its interaction with the noise associated with the estimation of the portfolio weights.

Differently, in the large scale example, each model is estimated for 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80 and 89 assets. In this second case, only daily data are available, so that the only possible covariance proxy is given by the cross product of the asset returns. As in the medium scale example, the assets are ordered alphabetically and, differently from the medium scale example, we will focus the attention on the model comparison tools which are less sensitive to the noise in the covariance proxy.

In both examples, we estimate the models with a 1-day rolling approach based on the last 2500 observations. In order to avoid dependence of the model comparison procedures on the mean return forecasts, these are always fixed at the sample mean[23]. All models for all problem dimensions are re-estimated daily, and are used to produce one-step-ahead forecasts. We consider two different out-of-sample evaluation periods. In the first, we focus on extreme market conditions and compare models for the period April 2008 – March 2009. This could be considered as a model stress test to determine if more highly parameterized models are preferred to simpler or naïve specifications as they are not exposed to parameter uncertainty and instability. The second forecast

[23] Simple diagnostic procedures on the mean returns support our choice given the extremely limited evidence of mean dynamics. As a consequence, the results reported are not biased by the misspecification of the mean dynamics.

evaluation period is for 2006, when the market was in a low volatility state and was trending upward. This second comparison allows testing of whether the model ranking might be affected by overall market conditions.

We stress that the empirical evaluations we report below might depend on the selection of equities, or on their ordering. However, appropriate evaluations of these elements are infeasible as they would require the estimation of all models on a large number of alternative asset orderings. We made a simple evaluation, not reported here, reverting the asset ordering on the large scale example. The results obtained do not support the possible effects of the equities selection as the model preferences are essentially equivalent to those reported below.[24] Differently, in the medium case example, when the number of assets is reduced, the introduction of a single equity might influence the result.

4.1 Medium scale example

If we consider pairwise direct model comparisons, the Amisano-Giacomini test highlights the poor performances of the covariance shrinking approach for both the out-of-sample periods. Furthermore, the EWMA is the second worst model, better than covariance shrinking but worse than most other models. Those results are only slightly influenced by the number of assets included in the evaluation. The Amisano-Giacomini test outcomes also suggest that the introduction of asymmetry in the variances or correlations induce some benefits only during the crisis period. In fact, in this second evaluation sample, models including asymmetry are generally preferred to the specifications without asymmetry. On the contrary, during a period of low market volatility, models with asymmetry provide forecasts equivalent to those obtained by models without asymmetry. Finally, dynamic conditional correlation specifications have statistically superior performances over constant conditional correlation models only during the crisis period.

If we perform pairwise direct comparisons using the Diebold-Mariano test, the results are also influenced by the choice of the covariance proxy. In this case, the preference

[24] A possible symptom of the effect of asset ordering would have been a clear change in the model preference for increasing problem dimension. Such an effect might be more evident when a small number of assets is used. This is one of the reasons which led us to focus on problem dimensions greater than or equal to 5 and 10 in the medium and large scale examples, respectively.

ordering across models is less evident, but we note that the impact of the noise in the covariance proxy is limited during the low market volatility period of 2006. In fact, the test outcomes are almost equivalent and show evidence of better performances of the dynamic and conditional correlation models compared with the other specifications. During the crisis period, the use of a noisy proxy influences the model ordering. Results obtained when the proxy is the cross product of realized returns are, for some model pairs, opposite to those provided by the test using the realized covariance.

In order to obtain a clearer picture of the model rankings, we move to the evaluation of the Model Confidence Set outcomes, which are reported in Tables 1 and 2. As the results for the two test statistics in (17) and (18) are substantially equivalent, we will refer in the following only to the test statistic in (17). With respect to the 2006 evaluation sample (Table 1), we note that the Amisano-Giacomini loss is basically excluding from the confidence set the EWMA, SHR and OGARCH models. Differently, under MSE loss, the use of a noisy covariance proxy makes most models equivalent (SHR exluded), while a realized covariance shows evidence of forecasting underperformance of the EWMA, SHR, BEKK models, OGARCH specifications and of GJR-cDCC. We also note that moving from 5 to 6 assets, the results are quite different, and seem to be stabilizing when the number of assets increases.

Moving to Table 2, the crisis evaluation period, the results are somewhat similar for the Amisano-Giacomini loss, but with a much clearer preference for DCC models, in particular for GJR-cDCC. On the contrary, using MSE loss, all the models are statistically equivalent when we use a noisy proxy, while some differences emerge (but only at the 5% confidence level) when using the realized covariance. In fact, at the 1% confidence level, both proxies lead to the same resulst. We might associate such an effect with the extreme volatility present in the market. Finally, for both sample periods, we also note that the introduction of asymmetry, either in the variances, or in the correlation or covariances, does not provide any improvements as the specifications with and without asymmetry are statistically equivalent.

We now shift to the indirect model evaluation, where we use both the Amisano-Giacomini loss and the MSE and QLIKE loss functions presented in Section 3. As this example has three loss functions, three portfolio strategies and two possible covariance proxies, we comment directly on the results of the Model Confidence Set approach. The evaluation of the outcomes of pairwise model comparisons does not lead to a clear picture of the model rankings.

We now examine Tables 3-5: in the first and second, we report the MSE Model Confidence Set p-values for two portfolio strategies (EW and GMV), the two choices of the covariance proxy, and the two evaluation periods; andTable 5 focuses on the QLIKE loss function. For MSE loss, we observe that the use of realized covariances plays a sensible role when the market volatility is not too high, as shown in Table 3. In fact, in the left panels all the models provide statistically equivalent forecasting performances, while in the right panel some preference across models clearly emerge.

As reported in Table 4, when the market is experiencing turbulence, all the models perform badly in forecasting the covariances, and the outcome is almost completely unaffected by the choice of the covariance proxy. Fortunately, such a result strongly depends on the loss function used. In fact, the QLIKE function shows evidence of a preference for some models, as reported in Table 5 for both evaluation periods.

If we combine the MCS outcomes over the two evaluation periods and the three loss functions[25], we can state the following: i) if we follow the Amisano-Giacomini approach, the preferred specifications are the ABEKK and GJR-OGARCH models; ii) there is an overall preference for asymmetryc CCC and DCC specifications (including GJR-CCC, GJR-DCC and GJR-ADCC) if we consider an equally weighted portfolio strategy; iii) when the portfolio weights are estimated, the GJR-OGARCH model is frequently included in the Model Confidence Set. Overall, the most successful models seem to be GJR-OGARCH and GJR-ADCC, quite possibly due to to their flexibility.

We further stress that the differences observed in Table 3 between the top right and bottom right panels, and those observed in Table 5 between the top and bottom panels, are influenced by the variability of portfolio weights. In fact, the top panels are using an equally weighted portfolio strategy, where the weights are calibrated, while the bottom panels refer to global minimum variance portfolios. These are estimated as a non-linear function of the covariance forecasts and, by construction, are time-varying. This induces an increase in the variability across models and over time, which influences the results. As a consequence, we believe that the comparisons across EW portfolios are entirely

[25] The unreported results for Amisano-Giacomini are available upon request.

relevant for an appropriate model evaluation as they are not affected by the estimation of portfolio weights.

What are the results if we compare models using the Engle and Colacito approaches? With respect to model accuracy (see equation (25)), most models are inaccurate, irrespective of the sample used, the portfolio strategy adopted, and of the covariance proxy employed. Such a result, as expected, is more evident during the crisis evaluation period.[26] Some minor evidence indicate a preference for the GJR-OGARCH model. Ranking models on the basis of the average expected portfolio variance often leads to inconsistent choices as the model underestimating the overall covariance is generally preferred (see Table 6). This happens, in particular, for the SHR approach during the financial crisis of 2008-2009 as the model does not react quickly to the increase in overall volatility.[27]

In summary, with respect to the possible ways of performing a model comparison, on the basis of our empirical results, we suggest the use of MCS with the Amisano-Giacomini loss as it does not depend on a covariance proxy. However, if a loss function based on a covariance proxy is preferred, we recommend the use of the QLIKE function. In both cases, the use of MCS of Hansen et al. (2003, 2010) is recommended.

Furthermore, considering the elements discussed in the introduction, the following conclusions emerge. First, the introduction of better covariance proxies has a relevant impact if we consider the MSE and QLIKE loss functions, in particular, during low volatility periods. Second, the performances of the naïve models are not really satisfactory, and this result does not depend on the problem size. In addition, the rankings across models do not seem to be time varying. Therefore, on the basis of our empirical results, we would conclude that the model specification and model estimation errors does not play a prominent role in the preference relations across the 13 models considered. Third, if we focus attention on the comparison between the DCC model of Engle (2002) and cDCC of Aielli (2008), our results favour the former. We link this outcome to the more complex estimation approach of the cDCC compared with that of DCC. Finally, in view of our empirical analyses, the results of Zumbach (2009) are only partially confirmed as the only covariance model which has relatively good

---

[26] The results are available upon request.

[27] Such an effect is the consequence of both the model structure as well as the size of the estimation window.

performance is OGARCH. Differently, the preference for CCC and DCC-type specifications is more evident.

4.2 Large scale example

Given the results of the medium scale example, we focus now only on the Model Confidence Set results (the pairwise comparisons do not provide unambiguous results). In addition, greater emphasis is given to the Amisano-Giacomini loss as it is not exposed to the noise in the covariance proxy.

In the direct comparison, the results of MCS for the Amisano-Giacomini loss are equivalent to those of the medium scale example: there is a preference for CCC and DCC-type models, in particular during the crisis (see Table 7), and the impact of asymmetry is limited. In addition, the MSE loss outcomes are consistent with the previous results using a noisy proxy for the crisis period, in that all models are equivalent. During 2006, there is a preference for the EWMA, GJR-ADCC and GJR-OGARCH models. The last two were included in the MCS of the medium scale example when the realized covariance was used as a covariance proxy. Even if the results cannot be verified (due to the absence of a realized covariance proxy for the 89 assets), the outcome partially confirms the previous finding of a mild preference in large asset cross sections for the GJR-ADCC and GJR-OGARCH specifications.

Moving to the indirect model evaluation, the Amisano-Giacomini loss on EW portfolios indicates that the BEKK and OGARCH specifications are preferred, consistently with what was observed in the medium scale example in both evaluation periods. GJR-OGARCH is the preferred model during the crisis evaluation period (Table 7), and the MCS includes also EWMA and OGARCH, but only at the 1% confidence level. The MSE and QLIKE loss functions have results that are similar to those observed with the noisy proxy in Section 4.1: most models are equivalent during 2008-2009, while there emerge some preference for the CCC, DCC and OGARCH models during 2006. In addition, the results change with respect to the portfolio strategy adopted, where the variability over time of the portfolio weights represents a potential source of noise.[28]

The outcomes of the Engle and Colacito (2006) approaches provide results that are consistent with those of the medium scale empirical examples: the test for model

[28] The unreported results are available upon request.

accuracy suggests that most models are inaccurate, but the results are biased by the noise in the covariance proxy used; the identification of the best model by means of portfolio variances often leads to a preference for naïve approaches, which underestimate the overall covariance (see Table 8).

Comparing the outcomes of the large scale example with those of the medium scale, we find confirmation of the poor performances of the naïve models and of the stability of the rankings across problem dimension (at least with the sample of assets and for the evaluation periods considered here). Therefore, the model estimation error seems not to have a prominent role in the model rankings.

5. Concluding Remarks

From an empirical perspective, Multivariate GARCH models suffer from the so-called curse of dimensionality. For this reason, several simple specifications are typically used, including the CCC, DCC, OGARCH and Scalar BEKK models. Alternatively, naïve methods could be used, such as EWMA or the Covariance Shrinking approach. However, few studies have considered a detailed out-of-sample comparison of these models. This paper has shed light on this topic, but the outcome is far from conclusive. By using alternative evaluation methods, including the direct and indirect approaches, pairwise and multivariate methodologies, realized covariance and noisy covariance proxies, and different out-of-sample evaluation periods, the results are mixed.

Some useful results emerge. The use of a realized covariance proxy is relevant as the rankings obtained with a noisy proxy can be quite different. This result complements the findings of Hansen and Lunde (2005, 2006) and Laurent et al. (2010). The rankings seem not to be greatly affected by the problem size: apart from some variability for the smallest problem dimensions considered, by increasing the number of assets the model ranking stabilizes as if the impact of model estimation and specification errors (which should be increasing with the problem dimension) are not affecting the rankings.

Furthermore, naïve approaches such as EWMA and covariance shrinking methods underperform compared with the dynamic models. Less common outcomes suggest that, during periods of high volatility, most models provide statistically equivalent results, while some preference is given to DCC-type specifications and GJR-OGARCH models. Across the methods considered, we highlight that the use of the MCS of Hansen et al. (2003, 2010) leads to results that are easiar to interpret, the Amisano-Giacomini (2007) approach is not influenced by the noise in the covariance proxy, while the QLIKE loss function seems to be able to detect some model preferences, even in periods of high volatility.

Overall, we do not find confirmation of the result of Zumbach (2009), which suggested a preference for covariance models. Furthermore, we provide evidence that naïve allocation strategies, such as EW, should be preferred as they are not influenced by the variability of the portfolio weights, which might have a role in the model rankings.

Finally, it should be emphasized that the main message from the empirical analysis is that there is no optimal model. The best model must be chosen with respect to a sample period and by using selection criteria that match the purpose of the analysis. It is clear that direct and indirect evaluations can provide markedly different results. This may be read as further confirmation of the truism that “all models are wrong, but some are more useful than others”, wherein usefulness may change over time and for different applications.

The fact that model rankings might change over time, and that alternative models are included in the MCS, might provide a reasonable data-driven input for a forecast combination of MGARCH specifications. A possible approach would then follow the ideas of Amendola and Storti (2009), who propose a methodology for forecast combination but restrict their attention to two standard MGARCH specifications.

Additional research on the topic is needed, and should focus on the methodological approaches for model comparison, on the rubustness of model rankings over different forecast horizons (longer that the one-day horizon we use), and on the impact of estimating the portfolio weights. Such tasks to be properly investigated would require the use of simulation-based approaches on a large cross-sectional dimension. We leave this computationally challenging topic to future research.

References

Aielli, G.P., 2006, Consistent estimation of large scale dynamic conditional correlations, Working paper n. 47, Department of Economics, Statistics, Mathematics and Sociology, University of Messina.

Alexander, C.O., 2001a, Orthogonal GARCH, in C.O. Alexander (ed.), Mastering Risk, vol. II, pp. 21–38. Prentice-Hall.

Alexander, C.O., 2001b, Market Models: A Guide to Financial Data Analysis, Chichester, U.K., Wiley.

Amendola, A., and Storti, G., 2009, Combination of multivariate volatility forecasts, SFB 649 Economic Risk, Discussion Paper 2009-007.

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

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

Andersen, T.G., Bollerslev, T. and Diebold, F.X., 2009, Parametric and non parametric volatility measurement, in L.P. Hansen and Y. Ait-Sahalia (eds.), Handbook of Financial Econometrics, Elsevier.

Asai, M., Caporin, M., and McAleer, M., 2009, Block structure multivariate stochastic volatility, Available at SSRN: http://ssrn.com/abstract=1524667.

Asai, M., M. McAleer and J. Yu (2006), Multivariate stochastic volatility: A review, Econometric Reviews, 25, 145-175.

Barndorff-Nielsen, O., and Shephard, N., 2004, Econometric analysis of realized covariation: High frequency covariance, regression and correlation in financial economics, Econometrica, 72, 885–925.

Barndorff-Nielsen, O., Hansen, P.R., Lunde, A., and Shephard, N., 2008, Multivariate realized kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-syncronous trading, CREATES Research Papers, 2008-63.

Bauwens, L., Laurent, S., and Rombouts, J.V.K., 2006, Multivariate GARCH models: A survey, Journal of Applied Econometrics, 21, 79-109.

Billio, M., Caporin, M. and Gobbo, M., 2006, Flexible dynamic conditional correlation multivariate GARCH for asset allocation, Applied Financial Economics Letters, 2, 123-130.

Bollerslev T., 1990, Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH approach, Review of Economic and Statistics, 72, 498-505.

Bonato, M., Caporin, M., and Ranaldo, A., 2009, Forecasting realized covariances with a block structure WAR model, Swiss National Bank Working Paper 2009-03.

Caporin, M. and McAleer, M., 2008, Scalar BEKK and indirect DCC, Journal of Forecasting, 27-6, 537-549.

Caporin, M. and McAleer, M., 2011, Do we really need both BEKK and DCC? A tale of two covariance models, Journal of Economic Surveys, forthcoming.

Caporin, M., and Paruolo, P., 2009, Structured multivariate volatility models, Available at SSRN: http://ssrn.com/abstract=1318639.

Cappiello, L., Engle, R.F., and Sheppard, K., 2006, Asymmetric dynamics in the correlations of global equity and bond returns, Journal of Financial Econometrics, 4, 537-572.

Chib, S., Omori, Y., and Asai, M., 2009, Multivariate Stochastic Volatility, in Andersen, T.G., Davis, R.A., Kreiß, J.P., and Mikosch, T. (eds.), Handbook of Financial Time Series, Springer.

Clements, A., Doolan, M., Hurn, S., and Becker, M., 2009, On the efficacy of techniques for evaluating multivariate volatility forecasts, NCER working paper series.

Comte, F. and Lieberman, O., 2003, Asymptotic theory for multivariate GARCH processes, Journal of Multivariate Analysis, 84, 61-84.

De Miguel, V., Garlappi, L., and Uppal, R., 2009, Optimal versus naïve diversification: How inefficient is the 1/N portfolio strategy?, Review of Financial Studies, 22, 1915-1953.

Diebold, F.X. and Mariano, R.S., 1995, Comparing predictive accuracy, Journal of Business and Economic Statistics, 13, 253-263.

Ding, Z. and Engle, R., 2001, Large scale conditional covariance modelling, estimation and testing, Academia Economic Papers, 29, 157-184.

Engle, R.F., 2002, Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models, Journal of Business and Economic Statistics, 20, 339-350.

Engle, R.F., and Colacito, R., 2006, Testing and valuing dynamic correlations for asset allocation, Journal of Business and Economic Statistics, 24, 238-253.

Engle, R.F., and Kelly, B., 2008, Dynamic equicorrelation, New York University Working Paper FIN-08-038.

Engle, R.F. and Kroner, K.F., 1995, Multivariate simultaneous generalized ARCH, Econometric Theory, 11, 122-150.

Engle, R.F., Ng, V.K., and Rothschild, M., 1990, Asset pricing with a factor ARCH covariance structure: empirical estimates for Treasury bills, Journal of Econometrics, 45, 213.238.

Engle, R.F., and Sheppard, K., 2001, Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH, Working Paper 2001-15, University of California at San Diego.

Engle, R.F., and Sheppard, K., 2008, Evaluating the specification of covariance models for large portfolios, available at www.kevinsheppard.net.

Engle, R.F., Shephard, N., and Sheppard, K., 2008, Fitting vast dimensional time-varying covariance models, Oxford Financial Research Centre, Financial Economics Working Paper n. 30.

Fan, Y., Pastorello, S., and Renault, E., 2007, Maximization by parts in extremum estimation, Mimeo, University of North Carolina in Chapel Hill.

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.

Franses, P.H., and Hafner, C.M., 2009, A generalized dynamic conditional correlation model: Simulation and application to many assets, Econometric Reviews, 28, 612-631.

Francq, C. and J.M. Zakoian, 2010, QML estimation of a class of multivariate GARCH models without moment conditions on the observed process, MPRA paper 20779 University Library of Munich, Germany, available at http://mpra.ub.uni-muenchen.de/20779/.

Glosten, L., Jagannathan, R., and Runkle, D., 1993, On the relationship between the expected value and the volatility of the nominal excess returns on stocks, Journal of Finance 48, 1779-1801.

Hafner, C.M., and Herwartz, H., 2008, Analytical quasi maximum likelihood inference in multivariate volatility models, Metrika, 67, 219-239.

Hafner, C.M., and Preminger, A., 2009, On asymptotic theory for multivariate GARCH models, Journal of Multivariate Analysis, 100, 2044-2054.

Hafner, C.M., and Reznikova, O., 2010, On the estimation of dynamic conditional correlation models, Université Catholique de Louvain, Institute of Statistics, Discussion Paper DS1006.

Hansen, P.R., 2005, A test for superior predictive ability, Journal of Business and Economic Statistics, 23-4, 365-380.

Hansen, P.R., and Lunde, A., 2005, A forecast comparison of volatility models: Does anything beat a GARCH(1,1)?, Journal of Applied Econometrics, 20, 873-889.

Hansen, P.R., and Lunde, A., 2006, Consistent ranking of volatility models, Journal of Econometrics, 131, 97-121.

Hansen, P.R., Lunde, A. and Nason, J.M., 2003, Choosing the best volatility models: The model confidence set approach, Oxford Bulletin of Economics and Statistics, 65, 839–861.

Hansen, P.R., Lunde, A. and Nason, J.M., 2010, The model confidence sets, Econometrica, forthcoming.

Jeantheau, T., 1998, Strong consistency of estimators for multivariate ARCH models, Econometric Theory, 14, 70-86.

Lanne, M., and Saikkonen, P., 2007, A multivariate generalized orthogonal factor GARCH model, Journal of Business and Economic Statistics, 25, 61-75.

Laurent, S., Rombouts, J.V.K., and Violante, F., 2009, On loss functions and ranking forecasting performances of multivariate GARCH models, CIRANO working paper 2009s-45.

Laurent, S., Rombouts, J.V.K., and Violante, F., 2010, On the forecasting accuracy of multivariate GARCH models, Journal of Applied Econometrics, forthcoming.

Ledoit, O., and Wolf, M., 2003, Improved estimation of the covariance matrix of stock returns with an application to portfolio selection, Journal of Empirical Finance, 10, 603-621.

Ledoit, O., and Wolf, M., 2004, Honey, I shrunk the sample covariance matrix, Journal of Portfolio Management, 2004, 110-119.

Ling, S. and McAleer, M., 2003, Asymptotic theory for a vector ARMA-GARCH model, Econometric Theory, 19, 278-308.

McAleer, M., 2005, Automated inference and learning in modeling finanical volatility, Econometric Theory, 21, 232-261.

McAleer, M., Chan, F., Hoti, S., and Lieberman, O., 2009, Generalized autoregressive conditional correlation, Econometric Theory, 24, 1554-1583.

McAleer M. and M. Medeiros, 2008, Realized volatility: A review, Econometric Reviews, 27, 10-45.

Mincer, J., and Zarnowitz, V., 1969, The evaluation of economic forecasts, in Mincer, J. (ed.) Economic Forecasts and Expectations, Columbia University Press.

Patton, A.J., 2010, Volatility forecast comparison using imperfect volatility proxies, Journal of Econometrics, forthcoming.

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., (eds.), Handbook of Financial Time Series, Springer.

Silvennoinen, A., and Terasvirta, T., 2009, Multivariate GARCH models, in Andersen, T.G., Davis, R.A., Kreiß, J.P., and Mikosch, T. (eds.), Handbook of Financial Time Series, Springer.

Tse, Y.K., and Tsui, A.K.C., 2002, A multivariate GARCH model with time-varying correlations, Journal of Business and Economic Statistics, 20, 351-362.

Van der Weide, R., 2002, GO.GARCH: A multivariate generalized orthogonal GARCH model, Journal of Applied Econometrics, 17, 549-564.

Voev, V., 2009, On the economic evaluation of volatility forecasts, CREATES WP 2009-56, School of Economics and Management, University of Aarhus.

Vrontos, I.D., Dellaportas, P., and Politis, D.N., 2003, A full-factor multivariate GARCH model, Econometrics Journal, 6, 312-334.

West, K.D., 1996, Asymptotic inference about predictive ability, Econometrica, 64, 1067-1084.

West, K.D., 2006, Forecast evaluation, in Elliott, G., Granger, C., Timmermann, A. (eds.) Handbook of Economic Forecasting, North Holland, Amsterdam.

White, H., 2000, A reality check for data snooping, Econometrica, 68-5, 1097-1126

Zumbach, G., 2009, The empirical properties of large covariance matrices, RiskMetrics Journal, 9-1 Winter 2009, 31-54.

Appendix A: List of equities included in the empirical analysis

The following is a list of the 89 companies whose stock total returns have been used in the empirical analysis of the paper for the large scale example:

3M, ABBOTT LABORATORIES, ALCOA, ALLSTATE, ALTRIA GROUP, AMER.ELEC.PWR., AMERICAN EXPRESS, AMGEN, APPLE, AT&T, AVON PRODUCTS, BAKER HUGHES, BANK OF AMERICA, BANK OF NEW YORK MELLON, BAXTER INTL., BOEING, BRISTOL MYERS SQUIBB, BURL.NTHN.SANTA FE C, CAMPBELL SOUP, CAPITAL ONE FINL., CATERPILLAR, CHEVRON, CISCO SYSTEMS, CITIGROUP, COCA COLA, COLGATE-PALM., COMCAST 'A', CONOCOPHILLIPS, COSTCO WHOLESALE, CVS CAREMARK, DELL, DOW CHEMICAL, E I DU PONT DE NEMOURS, EMC, ENTERGY, EXELON, EXXON MOBIL, FEDEX, FORD MOTOR, GENERAL DYNAMICS, GENERAL ELECTRIC, GILEAD SCIENCES, HALLIBURTON, HEWLETT-PACKARD, HJ HEINZ, HOME DEPOT, HONEYWELL INTL., INTEL, INTERNATIONAL BUS.MCHS., JOHNSON & JOHNSON, JP MORGAN CHASE & CO., LOCKHEED MARTIN, LOWE'S COMPANIES, MCDONALDS, MEDTRONIC, MERCK & CO., MICROSOFT, MORGAN STANLEY, NATIONAL OILWELL VARCO, NIKE 'B', NORFOLK SOUTHERN, OCCIDENTAL PTL., ORACLE, PEPSICO, PFIZER, PROCTER & GAMBLE, QUALCOMM, RAYTHEON 'B', REGIONS FINL.NEW, SARA LEE, SCHERING-PLOUGH, SCHLUMBERGER, SOUTHERN, SPRINT NEXTEL, TARGET, TEXAS INSTS., TIME WARNER, UNITED TECHNOLOGIES, UNITEDHEALTH GP., US BANCORP, VERIZON COMMUNICATIONS, WAL MART STORES, WALGREEN, WALT DISNEY, WELLS FARGO & CO, WEYERHAEUSER, WILLIAMS COS., WYETH, XEROX

The following is a list of the 15 companies whose total returns have been used in the empirical analysis of the paper for the medium scale example:

AT&T , BANK OF AMERICA, BOEING, CATERPILLAR, CITIGROUP, FEDEX, HONEYWELL INTL., HEWLETT-PACKARD, INTERNATIONAL BUS.MCHS., JP MORGAN CHASE & CO., PEPSICO, PROCTER & GAMBLE, TEXAS INSTS., TIME WARNER, WELLS FARGO & CO

Table 1: Model Confidence Set results for the 2006 evaluation period

Number of variables	5	6	7	8	9	10	11	12	13	14	15
Models – Loss function	**Amisano – Giacomini**										
EWMA	**0.16**	**0.02**	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SHR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CCC	**0.51**	**0.65**	**0.67**	**0.32**	**0.39**	**0.52**	**0.46**	**0.47**	**0.40**	**0.32**	**0.29**
DCC	**0.74**	**0.65**	**0.67**	**0.72**	**0.72**	**0.97**	**0.91**	**1.00**	**1.00**	**1.00**	**1.00**
cDCC	**0.51**	**0.65**	**0.71**	**0.32**	**0.39**	**0.52**	**0.46**	**0.47**	**0.40**	**0.32**	**0.29**
BEKK	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**0.47**	**0.40**	**0.32**	**0.29**
OGARCH	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GJR-CCC	**0.51**	**0.53**	**0.61**	**0.29**	**0.34**	**0.24**	**0.33**	**0.39**	**0.40**	**0.27**	**0.23**
GJR-DCC	**0.51**	**0.65**	**0.67**	**0.29**	**0.39**	**0.52**	**0.46**	**0.47**	**0.40**	**0.32**	**0.29**
GJR-ADCC	**0.51**	**0.65**	**0.67**	**0.29**	**0.39**	**0.52**	**0.46**	**0.47**	**0.40**	**0.32**	**0.29**
GJR-cDCC	**0.51**	**0.53**	**0.61**	**0.29**	**0.39**	**0.24**	**0.33**	**0.47**	**0.40**	**0.32**	**0.29**
ABEKK	**0.76**	**0.65**	**0.67**	**0.29**	**0.39**	**0.24**	**0.33**	**0.06**	**0.02**	**0.06**	**0.06**
GJR-OGARCH	0.00	**0.02**	**0.01**	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	Mean Squared Error (Diebold-Mariano) – Nosiy Proxy										
EWMA	**0.12**	**0.06**	**0.08**	**0.08**	**0.04**	**0.07**	**0.05**	**0.06**	**0.05**	**0.05**	**0.05**
SHR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CCC	**0.34**	**0.29**	**0.80**	**0.69**	**0.59**	**0.64**	**0.05**	**0.06**	**0.05**	**0.05**	**0.05**
DCC	**1.00**	**1.00**	**0.80**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**0.05**	**0.05**	**0.05**
cDCC	**0.34**	**0.29**	**1.00**	**0.69**	**0.59**	**0.64**	**0.56**	**0.34**	**0.05**	**0.05**	**0.05**
BEKK	**0.34**	**0.29**	**0.80**	**0.08**	**0.04**	**0.07**	**0.05**	**0.06**	**0.05**	**0.05**	**0.05**
OGARCH	**0.12**	**0.06**	**0.08**	**0.08**	**0.04**	**0.07**	**0.05**	**0.03**	**0.05**	**0.05**	**0.04**
GJR-CCC	**0.12**	**0.06**	**0.08**	**0.08**	**0.04**	**0.07**	**0.05**	**0.06**	**0.05**	**0.05**	**0.05**
GJR-DCC	**0.12**	**0.06**	**0.08**	**0.08**	**0.04**	**0.07**	**0.05**	**0.06**	**0.05**	**0.05**	**0.05**
GJR-ADCC	**0.34**	**0.29**	**0.80**	**0.69**	**0.93**	**0.98**	**0.90**	**0.95**	**1.00**	**1.00**	**1.00**
GJR-cDCC	**0.12**	**0.06**	**0.08**	**0.08**	**0.04**	**0.07**	**0.05**	**0.06**	**0.05**	**0.05**	**0.05**
ABEKK	**0.34**	**0.06**	**0.08**	**0.08**	**0.04**	**0.07**	**0.05**	**0.06**	**0.05**	**0.05**	**0.05**
GJR-OGARCH	**0.12**	**0.06**	**0.08**	**0.08**	**0.04**	**0.07**	**0.05**	**0.06**	**0.05**	**0.05**	**0.05**
	Mean Squared Error (Diebold-Mariano) – Realized Covariance										
EWMA	0.00	**0.03**	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SHR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CCC	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
DCC	**0.50**	**0.31**	**0.59**	**0.50**	**0.29**	**0.37**	**0.50**	**0.40**	**0.64**	**0.62**	**0.75**
cDCC	0.00	**0.03**	**0.59**	**0.50**	**0.29**	**0.35**	**0.49**	**0.40**	**0.64**	**0.62**	**0.75**
BEKK	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OGARCH	0.00	**0.03**	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GJR-CCC	0.00	**0.03**	**0.45**	**0.50**	**0.38**	**0.41**	**0.68**	**0.48**	**0.64**	**0.62**	**0.75**
GJR-DCC	0.00	**0.03**	**0.07**	**0.05**	**0.12**	**0.18**	**0.17**	**0.12**	**0.22**	**0.09**	**0.04**
GJR-ADCC	0.00	**0.03**	**0.30**	**0.30**	**0.29**	**0.35**	**0.49**	**0.40**	**0.64**	**0.62**	**0.73**
GJR-cDCC	0.00	**0.03**	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ABEKK	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GJR-OGARCH	**0.50**	**0.03**	0.01	0.00	**0.29**	**0.37**	**0.99**	0.00	0.00	**0.09**	**0.92**

Note: Bold shaded p-values denote models included in the confidence set for each of the problem dimensions reported in the first row at the 1% confidence level.

Table 2: Model Confidence Set results for the crisis period (2008-2009)

Number of variables	5	6	7	8	9	10	11	12	13	14	15
Model – Loss function						**Amisano-Giacomini**					
EWMA	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SHR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CCC	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DCC	**0.32**	**0.12**	**0.18**	**0.19**	**0.17**	**0.09**	**0.04**	**0.04**	**0.03**	**0.04**	**0.09**
cDCC	**0.13**	**0.11**	**0.18**	**0.19**	**0.17**	**0.09**	**0.04**	**0.04**	**0.03**	**0.04**	**0.09**
BEKK	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OGARCH	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GJR-CCC	**0.02**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00
GJR-DCC	**0.50**	**0.34**	**0.39**	**0.19**	**0.17**	**0.09**	**0.04**	**0.04**	**0.03**	**0.04**	**0.09**
GJR-ADCC	**0.50**	**0.34**	**0.50**	**0.71**	**0.74**	**0.28**	**0.79**	**1.00**	**1.00**	**0.79**	**0.51**
GJR-cDCC	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**0.98**	**0.77**	**1.00**	**1.00**
ABEKK	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GJR-OGARCH	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
						Mean Squared Error (Diebold-Mariano) – Nosiy Proxy					
EWMA	**0.31**	**0.26**	**0.21**	**0.22**	**0.31**	**0.28**	**0.24**	**0.24**	**0.20**	**0.10**	**0.09**
SHR	**0.05**	**0.04**	**0.04**	**0.05**	**0.04**	**0.04**	**0.04**	**0.03**	**0.04**	**0.03**	**0.03**
CCC	**0.31**	**0.26**	**0.21**	**0.22**	**0.27**	**0.28**	**0.24**	**0.21**	**0.20**	**0.10**	**0.09**
DCC	**0.31**	**0.26**	**0.21**	**0.22**	**0.27**	**0.28**	**0.24**	**0.21**	**0.20**	**0.10**	**0.09**
cDCC	**0.31**	**0.26**	**0.21**	**0.22**	**0.27**	**0.28**	**0.24**	**0.21**	**0.20**	**0.10**	**0.09**
BEKK	**0.31**	**0.26**	**0.21**	**0.22**	**0.27**	**0.28**	**0.24**	**0.21**	**0.20**	**0.10**	**0.09**
OGARCH	**0.25**	**0.25**	**0.21**	**0.19**	**0.24**	**0.28**	**0.24**	**0.21**	**0.20**	**0.10**	**0.09**
GJR-CCC	**0.31**	**0.26**	**0.21**	**0.22**	**0.27**	**0.28**	**0.24**	**0.21**	**0.20**	**0.10**	**0.09**
GJR-DCC	**0.31**	**0.26**	**0.21**	**0.22**	**0.31**	**0.28**	**0.24**	**0.24**	**0.20**	**0.10**	**0.09**
GJR-ADCC	**1.00**	**1.00**	**1.00**	**1.00**	**0.90**	**0.28**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
GJR-cDCC	**0.54**	**0.64**	**0.90**	**0.38**	**1.00**	**1.00**	**0.46**	**0.81**	**0.37**	**0.47**	**0.44**
ABEKK	**0.31**	**0.26**	**0.21**	**0.22**	**0.27**	**0.28**	**0.24**	**0.21**	**0.20**	**0.10**	**0.09**
GJR-OGARCH	**0.31**	**0.26**	**0.21**	**0.22**	**0.25**	**0.28**	**0.24**	**0.21**	**0.20**	**0.10**	**0.09**
						Mean Squared Error (Diebold-Mariano) – Realized Covariance					
EWMA	0.01	**0.02**	**0.02**	**0.01**	**0.02**	**0.03**	**0.03**	**0.03**	**0.03**	**0.04**	**0.04**
SHR	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
CCC	**0.01**	**0.02**	**0.02**	**0.01**	**0.02**	**0.03**	**0.03**	**0.03**	**0.03**	**0.04**	**0.04**
DCC	0.01	**0.02**	**0.02**	**0.01**	**0.02**	**0.03**	**0.03**	**0.03**	**0.03**	**0.04**	**0.04**
cDCC	0.01	**0.02**	**0.02**	**0.01**	**0.02**	**0.03**	**0.03**	**0.03**	**0.03**	**0.04**	**0.04**
BEKK	**0.02**	**0.04**	**0.07**	**0.06**	**0.23**	**0.21**	**0.23**	**0.37**	**0.54**	**0.25**	**0.36**
OGARCH	**0.02**	**0.04**	**0.07**	**0.78**	**0.24**	**0.68**	**0.85**	**0.37**	**0.54**	**0.25**	**0.36**
GJR-CCC	**0.02**	**0.03**	**0.02**	**0.01**	**0.02**	**0.03**	**0.03**	**0.03**	**0.03**	**0.04**	**0.04**
GJR-DCC	**0.01**	**0.03**	**0.02**	**0.01**	**0.02**	**0.03**	**0.03**	**0.03**	**0.03**	**0.04**	**0.04**
GJR-ADCC	**0.01**	**0.03**	**0.02**	**0.02**	**0.02**	**0.03**	**0.03**	**0.03**	**0.03**	**0.04**	**0.04**
GJR-cDCC	0.01	**0.02**	**0.02**	**0.01**	**0.02**	**0.03**	**0.03**	**0.03**	**0.03**	**0.04**	**0.04**
ABEKK	**0.02**	**0.04**	**0.05**	**0.03**	**0.06**	**0.06**	**0.07**	**0.06**	**0.09**	**0.08**	**0.10**
GJR-OGARCH	**0.02**	**0.04**	**0.07**	**0.78**	**0.24**	**0.40**	**0.46**	**0.37**	**0.54**	**0.08**	**0.10**

Note: Bold shaded p-values denote models included in the confidence set for each of the problem dimensions reported in the first row at the 1% confidence level.

Table 3: Model Confidence Set results for MSE loss under indirect model comparison for the 2006 evaluation period

	Number of variables	5	6	7	8	9	10	11	12	13	14	15	5	6	7	8	9	10	11	12	13	14	15
		Covariance proxy: returns cross-product											**Covariance proxy: realized covariance**										
Equally weighted porfolio	**EWMA**	**0.14**	**0.04**	**0.09**	**0.10**	**0.08**	**0.10**	**0.06**	**0.03**	**0.01**	**0.03**	**0.04**	0.00	**0.02**	**0.01**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	SHR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	CCC	**0.14**	**0.04**	**0.09**	**0.10**	**0.08**	**0.10**	**0.06**	**0.03**	**0.01**	**0.03**	**0.04**	**0.53**	**0.24**	**0.13**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	DCC	**0.91**	**0.04**	**0.09**	**0.10**	**0.08**	**0.10**	**0.06**	**0.03**	**0.01**	**0.03**	**0.03**	0.00	**0.02**	**0.01**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	cDCC	**0.91**	**0.04**	**0.49**	**0.10**	**0.08**	**0.10**	**0.06**	**0.02**	0.01	**0.02**	**0.03**	0.00	**0.02**	**0.01**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	BEKK	**0.14**	**0.04**	**0.09**	**0.10**	**0.08**	**0.10**	**0.06**	**0.02**	0.01	**0.02**	**0.03**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	OGARCH	**0.91**	**0.04**	**0.09**	**0.10**	**0.08**	**0.10**	**0.08**	**0.03**	**0.03**	**0.03**	**0.03**	0.00	**0.02**	**0.01**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	GJR-CCC	**0.14**	**0.04**	**0.59**	**0.62**	**0.46**	**0.48**	**0.08**	**0.41**	**0.42**	**0.48**	**0.54**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**0.98**
	GJR-DCC	**0.14**	**0.04**	**0.09**	**0.10**	**0.08**	**0.10**	**0.06**	**0.03**	**0.03**	**0.03**	**0.05**	**0.33**	**0.24**	**0.13**	**0.03**	**0.03**	**0.03**	0.00	**0.01**	**0.02**	**0.02**	**0.05**
	GJR-ADCC	**0.91**	**0.78**	**0.59**	**0.62**	**0.46**	**0.48**	**0.23**	**0.41**	**0.42**	**0.48**	**0.54**	**0.53**	**0.24**	**0.61**	**0.60**	**0.62**	**0.57**	**0.65**	**0.61**	**0.51**	**0.58**	**1.00**
	GJR-cDCC	**0.14**	**0.04**	**0.09**	**0.10**	**0.08**	**0.10**	**0.06**	**0.03**	**0.03**	**0.03**	**0.04**	0.00	**0.02**	**0.01**	0.00	0.00	0.01	0.00	0.00	0.01	0.00	**0.01**
	ABEKK	**0.14**	**0.04**	**0.09**	**0.10**	**0.08**	**0.10**	**0.06**	**0.02**	0.01	**0.02**	**0.03**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	GJR-OGARCH	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.01**
		Covariance proxy: returns cross-product											**Covariance proxy: realized covariance**										
Global Minimum Vairance portfolio	**EWMA**	**0.14**	**0.43**	**0.51**	**0.73**	**0.67**	**0.67**	**0.16**	**0.42**	**0.42**	**0.15**	**0.22**	**0.14**	**0.16**	**0.42**	**0.23**	**0.09**	**0.05**	**0.04**	**0.10**	**0.07**	0.00	0.00
	SHR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	CCC	**0.14**	**0.12**	**0.19**	**0.60**	**0.67**	**0.67**	**0.67**	**1.00**	**1.00**	**0.85**	**0.69**	**0.47**	**0.65**	**0.58**	**0.47**	**0.67**	**0.29**	**0.35**	**0.86**	**0.92**	0.00	0.00
	DCC	**0.34**	**0.12**	**0.51**	**0.85**	**1.00**	**1.00**	**0.98**	**0.97**	**0.96**	**0.85**	**0.69**	**0.14**	**0.27**	**0.47**	**0.40**	**0.17**	**0.21**	**0.14**	**0.40**	**0.45**	0.00	0.00
	cDCC	**0.14**	**0.12**	**0.51**	**0.85**	**0.96**	**1.00**	**1.00**	**0.97**	**0.96**	**0.85**	**0.37**	**0.14**	**0.27**	**0.58**	**0.40**	**0.09**	**0.21**	**0.13**	**0.32**	**0.45**	0.00	0.00
	BEKK	**1.00**	**1.00**	**1.00**	**1.00**	**0.96**	**1.00**	**0.67**	**0.97**	**0.96**	**0.85**	**0.70**	0.00	0.00	0.00	**0.40**	**0.67**	**0.45**	**0.35**	**0.86**	**0.92**	**0.13**	0.00
	OGARCH	**0.14**	**0.12**	**0.19**	**0.17**	**0.18**	**0.23**	**0.16**	**0.27**	**0.28**	**0.15**	**0.22**	0.00	0.00	0.00	**0.12**	**0.09**	**0.05**	**0.14**	**0.10**	**0.07**	**0.27**	**0.16**
	GJR-CCC	**0.14**	**0.12**	**0.19**	**0.45**	**0.67**	**0.67**	**0.16**	**0.42**	**0.42**	**0.25**	**0.37**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**0.29**	**0.35**	**0.86**	**0.92**	0.00	0.00
	GJR-DCC	**0.92**	**0.95**	**0.65**	**0.74**	**0.96**	**0.67**	**0.16**	**0.42**	**0.42**	**0.15**	**0.37**	**0.31**	**0.27**	**0.58**	**0.40**	**0.25**	**0.21**	**0.13**	**0.32**	**0.35**	0.00	0.00
	GJR-ADCC	**0.37**	**0.43**	**0.51**	**0.45**	**0.67**	**0.67**	**0.16**	**0.42**	**0.42**	**0.15**	**0.22**	**0.31**	**0.22**	**0.47**	**0.40**	**0.25**	**0.21**	**0.13**	**0.32**	**0.35**	0.00	0.00
	GJR-cDCC	**0.37**	**0.43**	**0.51**	**0.45**	**0.96**	**0.67**	**0.16**	**0.42**	**0.42**	**0.15**	**0.22**	**0.31**	**0.27**	**0.58**	**0.40**	**0.09**	**0.05**	**0.04**	**0.10**	**0.07**	0.00	0.00
	ABEKK	**0.92**	**0.95**	**0.65**	**0.74**	**0.96**	**1.00**	**0.64**	**0.97**	**0.96**	**1.00**	**1.00**	0.00	0.00	0.00	**0.12**	**0.09**	**1.00**	**1.00**	**1.00**	**1.00**	**0.97**	**0.07**
	GJR-OGARCH	**0.14**	**0.12**	**0.19**	**0.20**	**0.29**	**0.35**	**0.16**	**0.27**	**0.34**	**0.15**	**0.22**	0.00	0.00	0.00	**0.40**	**0.67**	**0.05**	**0.35**	**0.10**	**0.07**	**1.00**	**1.00**

Note: Bold shaded p-values denote models included in the confidence set for each of the problem dimensions reported in the first row at the 1% confidence level.

Table 4: Model Confidence Set results for MSE loss under indirect model comparison for the crisis evaluation period

	Number of variables	5	6	7	8	9	10	11	12	13	14	15	5	6	7	8	9	10	11	12	13	14	15
		Covariance proxy: returns cross-product											Covariance proxy: realized covariance										
Equally weighted porfolio	EWMA	0.19	0.28	0.25	0.17	0.30	0.41	0.24	0.34	0.40	0.50	0.52	0.02	0.08	0.17	0.18	0.08	0.25	0.08	0.03	0.08	0.08	0.06
	SHR	0.04	0.03	0.03	0.02	0.02	0.03	0.02	0.02	0.03	0.02	0.02	1.00	0.29	0.17	0.06	0.08	0.11	0.08	0.03	0.08	0.08	0.06
	CCC	0.30	0.28	0.23	0.14	0.26	0.26	0.21	0.20	0.18	0.15	0.17	0.07	0.78	0.54	0.18	0.35	0.42	0.35	0.34	0.08	0.36	0.33
	DCC	0.30	0.28	0.25	0.17	0.35	0.41	0.24	0.34	0.27	0.22	0.23	0.02	0.10	0.23	0.18	0.08	0.42	0.08	0.03	0.08	0.08	0.06
	cDCC	0.30	0.28	0.23	0.14	0.30	0.35	0.21	0.20	0.18	0.22	0.23	0.02	0.10	0.17	0.18	0.08	0.25	0.08	0.03	0.08	0.08	0.06
	BEKK	0.10	0.10	0.10	0.14	0.11	0.14	0.17	0.20	0.18	0.13	0.17	0.07	0.78	0.37	0.18	0.08	0.25	0.08	0.03	0.08	0.08	0.06
	OGARCH	0.30	0.28	0.25	0.17	0.35	0.41	0.24	0.34	0.40	0.50	0.52	0.02	0.08	0.17	0.18	0.08	0.25	0.08	0.03	0.08	0.08	0.06
	GJR-CCC	0.30	0.28	0.25	0.17	0.27	0.26	0.21	0.20	0.18	0.15	0.17	0.07	1.00	1.00	0.18	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	GJR-DCC	0.30	0.28	0.25	0.17	0.35	0.41	0.24	0.34	0.27	0.22	0.23	0.02	0.10	0.23	0.77	0.35	0.61	0.86	0.34	0.75	0.36	0.33
	GJR-ADCC	0.41	0.32	0.25	0.17	0.45	0.41	0.24	0.34	0.40	0.50	0.52	0.02	0.10	0.23	1.00	0.35	0.42	0.86	0.66	0.92	0.36	0.33
	GJR-cDCC	0.41	0.32	0.25	0.17	0.45	0.41	0.24	0.34	0.40	0.49	0.45	0.02	0.08	0.17	0.18	0.08	0.42	0.08	0.03	0.08	0.08	0.06
	ABEKK	0.09	0.08	0.08	0.10	0.10	0.11	0.13	0.17	0.18	0.13	0.17	0.07	0.29	0.25	0.18	0.08	0.25	0.08	0.03	0.08	0.08	0.06
	GJR-OGARCH	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.02	0.08	0.17	0.18	0.08	0.25	0.08	0.03	0.08	0.08	0.06
		Covariance proxy: returns cross-product											Covariance proxy: realized covariance										
Global Minimum Vairance portfolio	EWMA	0.58	0.58	0.67	0.57	0.61	0.60	0.70	0.39	0.43	0.37	0.39	0.04	0.28	0.44	0.27	0.20	0.08	0.18	0.08	0.08	0.07	0.06
	SHR	0.01	0.00	0.01	0.01	0.01	0.07	0.15	0.26	0.15	0.05	0.10	0.04	1.00	0.22	0.27	0.20	0.08	0.09	0.08	0.08	0.20	0.06
	CCC	0.58	0.81	0.93	0.57	0.61	0.56	0.15	0.31	0.36	0.23	0.39	1.00	0.28	0.12	0.27	0.29	0.25	0.18	0.43	0.31	0.37	0.06
	DCC	0.58	0.81	0.80	0.49	0.61	0.51	0.15	0.31	0.36	0.33	0.39	0.07	0.28	0.17	0.27	0.27	0.28	0.22	0.43	0.31	0.37	0.06
	cDCC	0.58	0.77	0.70	0.57	0.61	0.56	0.17	0.31	0.43	0.37	0.58	0.04	0.28	0.22	0.27	0.29	0.28	0.22	0.43	0.31	0.37	0.06
	BEKK	0.58	0.77	0.68	0.57	0.61	0.60	1.00	1.00	1.00	1.00	1.00	0.02	0.50	1.00	0.27	0.27	0.25	0.14	0.08	0.06	0.20	0.03
	OGARCH	0.58	0.58	0.48	0.49	0.29	0.60	0.17	0.39	0.15	0.18	0.58	0.02	0.05	0.03	0.22	0.20	0.20	0.18	0.42	0.17	0.20	0.06
	GJR-CCC	1.00	1.00	1.00	0.57	0.61	0.56	0.17	0.31	0.36	0.31	0.39	0.88	0.51	0.23	1.00	1.00	0.25	0.18	0.43	0.31	0.37	0.96
	GJR-DCC	0.58	0.81	0.70	0.57	0.61	0.56	0.15	0.31	0.43	0.37	0.39	0.07	0.51	0.23	0.40	0.29	0.28	0.22	0.43	0.31	0.37	0.32
	GJR-ADCC	0.58	0.58	0.68	0.49	0.61	0.51	0.15	0.31	0.43	0.33	0.39	0.04	0.51	0.44	0.28	0.29	0.28	0.92	1.00	1.00	1.00	1.00
	GJR-cDCC	0.58	0.77	0.70	0.49	0.61	0.56	0.15	0.31	0.43	0.37	0.39	0.04	0.28	0.22	0.28	0.37	1.00	1.00	0.57	0.31	0.37	0.06
	ABEKK	0.58	0.58	0.70	1.00	1.00	1.00	0.97	0.44	0.43	0.37	0.58	0.02	0.28	0.44	0.27	0.27	0.20	0.13	0.08	0.05	0.20	0.04
	GJR-OGARCH	0.50	0.00	0.01	0.26	0.23	0.56	0.15	0.31	0.20	0.23	0.58	0.02	0.06	0.06	0.27	0.20	0.08	0.13	0.08	0.16	0.20	0.06

Note: Bold shaded p-values denote models included in the confidence set for each of the problem dimensions reported in the first row at the 1% confidence level.

Table 5: Model Confidence Set results for QLIKE loss under indirect model comparison using the realized covariance

	Number of variables	5	6	7	8	9	10	11	12	13	14	15	5	6	7	8	9	10	11	12	13	14	15
		2006 evaluation period											**Crisis evaluation period**										
Equally weighted portfolio	**EWMA**	0.01	**0.04**	**0.02**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	**0.03**	**0.01**	0.01	**0.01**	0.00	**0.01**
	SHR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.02**	0.00	0.00	0.01	0.00	**0.03**	**0.01**	0.01	**0.01**	0.00	**0.01**
	CCC	0.01	**0.04**	**0.02**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	**0.23**	**0.10**	**0.22**	0.01	0.00	**0.17**	**0.01**	0.01	**0.01**	0.00	**0.01**
	DCC	0.01	**0.04**	**0.02**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	**0.12**	**0.01**	0.01	**0.01**	0.00	**0.01**
	cDCC	0.01	**0.03**	**0.02**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	**0.03**	**0.01**	0.01	**0.01**	0.00	**0.01**
	BEKK	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.02**	0.00	0.00	0.01	0.00	**0.03**	**0.01**	0.01	**0.01**	0.00	**0.01**
	OGARCH	0.01	**0.03**	**0.02**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	**0.17**	**0.01**	0.01	**0.01**	0.00	**0.01**
	GJR-CCC	**0.86**	**0.86**	**0.71**	**0.73**	**0.62**	**0.92**	**0.80**	**0.97**	**0.94**	**0.97**	**0.73**	**1.00**	**1.00**	**1.00**	0.01	**0.94**	**0.86**	**0.44**	**0.34**	**0.01**	**1.00**	**0.41**
	GJR-DCC	**0.10**	**0.13**	**0.02**	0.01	**0.01**	**0.01**	0.00	0.01	0.00	0.01	**0.04**	0.00	0.00	**0.22**	**0.23**	**0.33**	**0.86**	**0.44**	**0.34**	**0.22**	**0.60**	**0.41**
	GJR-ADCC	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**0.80**	**1.00**	**1.00**	**0.97**	**0.73**	0.00	0.00	**0.22**	**0.28**	**1.00**	**0.77**	**0.76**	**1.00**	**1.00**	**0.83**	**1.00**
	GJR-cDCC	0.01	**0.04**	**0.02**	0.00	0.00	0.01	0.00	0.00	0.00	0.00	**0.01**	0.00	0.00	0.00	0.01	0.00	**0.77**	**0.01**	0.01	**0.01**	0.00	**0.01**
	ABEKK	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	**0.03**	**0.01**	0.01	**0.01**	0.00	**0.01**
	GJR-OGARCH	0.01	**0.03**	**0.02**	0.00	0.00	**0.92**	**1.00**	**0.97**	0.00	**1.00**	**1.00**	0.00	0.00	0.00	**1.00**	0.00	**1.00**	**1.00**	0.01	**0.01**	0.00	**0.01**
Global Minimum Vairance portfolio	**EWMA**	**0.11**	**0.28**	**0.07**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.01**	0.00	0.00	0.00	**0.26**	0.01	0.00	0.00	0.00	0.00	0.00
	SHR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	0.00	**0.01**	0.01	**0.01**	**0.01**	0.01	**0.04**	**0.05**
	CCC	**0.11**	**0.28**	**0.07**	0.00	0.00	**0.04**	0.00	**0.36**	0.00	0.00	0.00	**0.49**	0.00	0.00	0.00	**0.32**	**0.02**	**0.01**	**0.02**	**0.01**	**0.05**	**0.05**
	DCC	**0.08**	**0.05**	**0.07**	0.00	0.00	**0.02**	0.00	0.00	0.00	0.00	0.00	**0.01**	0.00	0.00	0.00	**0.33**	**0.02**	**0.01**	**0.02**	**0.01**	**0.05**	**0.05**
	cDCC	**0.06**	**0.05**	**0.07**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.01**	0.00	0.00	0.00	**0.33**	**0.02**	**0.22**	**0.02**	**0.01**	**0.05**	**0.05**
	BEKK	0.00	0.00	0.00	0.00	0.00	**0.54**	0.00	**0.56**	**0.70**	0.00	0.00	0.00	0.00	0.00	**0.03**	**0.33**	**0.02**	**0.01**	**0.02**	0.01	**0.04**	0.00
	OGARCH	**0.06**	**0.05**	**0.07**	**0.03**	0.00	**0.04**	0.00	0.00	0.00	0.00	**0.01**	0.00	0.00	0.00	**0.25**	**0.77**	**0.10**	**1.00**	**0.38**	**0.01**	**0.05**	**0.05**
	GJR-CCC	**1.00**	**1.00**	**1.00**	0.01	0.00	**0.04**	0.00	0.00	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	**0.33**	**0.02**	**0.01**	**0.02**	**0.01**	**0.04**	**0.05**
	GJR-DCC	**0.11**	**0.28**	**0.07**	0.00	0.00	**0.03**	0.00	0.00	0.00	0.00	0.00	**0.80**	0.00	0.00	0.00	**0.33**	**0.02**	**0.01**	**0.02**	**0.01**	**0.05**	**0.05**
	GJR-ADCC	**0.32**	**0.28**	**0.22**	0.00	0.00	**0.03**	0.00	0.00	0.00	0.00	0.00	**0.80**	0.00	0.00	0.00	**0.35**	**0.02**	**0.22**	**0.38**	**0.01**	**0.05**	**0.05**
	GJR-cDCC	**0.11**	**0.28**	**0.07**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.01**	0.00	0.00	0.00	**0.35**	**0.10**	**0.22**	**0.36**	**0.01**	**0.05**	**0.05**
	ABEKK	0.00	0.00	0.00	0.00	0.00	**1.00**	**0.33**	**1.00**	**0.73**	**0.80**	**0.01**	0.01	0.00	0.00	0.00	**0.33**	**0.02**	**0.01**	**0.02**	0.01	**0.05**	0.00
	GJR-OGARCH	**0.08**	**0.05**	**0.07**	**1.00**	**1.00**	**0.54**	**1.00**	**0.48**	**1.00**	**1.00**	**0.58**	0.00	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**0.89**	**1.00**	**1.00**	**1.00**	**1.00**

Note: Bold shaded p-values denote models included in the confidence set for each of the problem dimensions reported in the first row at the 1% confidence level.

Table 6: Model preference using Engle and Colacito (2006) approach in the medium scale example

	2006			**2008-2009**		
	EW	**GMV**	**GMVB**	**EW**	**GMV**	**GMVB**
EWMA	8.00	2.18	6.91	11.00	7.64	11.00
SHR	13.00	13.00	13.00	1.00	1.00	1.00
CCC	4.64	6.45	5.82	3.55	2.00	2.18
DCC	6.73	7.55	6.91	7.00	5.91	4.27
cDCC	6.27	6.91	6.36	7.36	7.09	5.64
BEKK	9.45	9.45	9.45	2.91	6.45	9.00
OGARCH	10.73	11.00	11.18	12.00	12.00	12.00
GJR-CCC	1.00	2.27	1.55	2.55	3.00	2.82
GJR-DCC	3.09	3.45	2.27	6.91	7.27	5.55
GJR-ADCC	2.09	4.09	2.55	8.73	8.27	6.91
GJR-cDCC	4.18	3.09	3.64	10.00	9.45	7.64
ABEKK	12.00	11.64	11.45	5.00	7.91	10.00
GJR-OGARCH	9.82	9.91	9.91	13.00	13.00	13.00

Note: The table reports the average ranking of each row model over the different cross-sectional dimensions we are considering. For each cross-sectional dimension the models are ranked on the basis of each portfolio strategy out-of-sample average portfolio expected variance. Shaded areas denote the best three models for each column.

Table 8: Model preference using Engle and Colacito (2006) approach in the large scale example

	2006			**2008-2009**		
	EW	GMV	GMVB	EW	GMV	GMVB
EWMA	1.00	1.00	1.00	11.00	1.92	11.00
SHR	13.00	13.00	13.00	1.00	3.08	1.00
CCC	10.15	6.08	6.08	2.92	2.77	2.00
DCC	7.77	7.69	7.75	5.69	6.77	4.00
cDCC	8.54	7.46	7.42	6.69	7.15	5.00
BEKK	10.46	11.00	11.00	3.62	5.31	9.15
OGARCH	3.00	10.00	10.00	12.00	12.00	12.00
GJR-CCC	8.00	2.15	2.08	4.23	4.23	3.00
GJR-DCC	5.62	4.85	4.92	7.85	9.69	6.00
GJR-ADCC	4.31	3.15	3.25	9.85	8.69	8.00
GJR-cDCC	5.62	4.15	4.08	8.85	10.23	7.00
ABEKK	11.54	12.00	12.00	4.31	6.15	9.85
GJR-OGARCH	2.00	8.46	8.42	13.00	13.00	13.00

Note: The table reports the average ranking of each row model over the different cross-sectional dimensions we are considering. For each cross-sectional dimension the models are ranked on the basis of each portfolio strategy out-of-sample average portfolio expected variance. Shaded areas denote the best three models for each column.

Table 7: Model Confidence Set results for the large scale example – Crisis evaluation period

Number of variables	10	15	20	25	30	35	40	45	50	60	70	80	89
Model Loss-function	**Amisano-Giacomini – Direct model comparison**												
EWMA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SHR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CCC	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DCC	**0.20**	**0.11**	**0.02**	**0.02**	**0.01**	**0.01**	**0.04**	**0.03**	**0.02**	**0.04**	**0.05**	**0.04**	**0.01**
cDCC	**0.20**	**1.00**	**1.00**	**1.00**	**0.88**	**0.01**	**0.04**	**0.03**	**0.06**	**0.05**	**0.05**	**0.04**	**0.01**
BEKK	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OGARCH	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GJR-CCC	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GJR-DCC	**0.20**	**0.11**	**0.02**	**0.02**	**0.01**	**0.01**	**0.04**	**0.03**	**0.04**	**0.05**	**0.05**	**0.04**	**0.01**
GJR-ADCC	**1.00**	**0.11**	**0.02**	**0.92**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
GJR-cDCC	**0.20**	**0.11**	**0.02**	**0.02**	**0.03**	**0.01**	**0.04**	**0.03**	**0.06**	**0.05**	**0.05**	**0.04**	**0.01**
ABEKK	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GJR-OGARCH	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Amisano-Giacomini – Indirect model comparison												
EWMA	**0.04**	**0.07**	**0.07**	**0.03**	**0.04**	**0.03**	**0.03**	**0.02**	**0.02**	**0.03**	**0.03**	**0.03**	**0.03**
SHR	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CCC	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DCC	**0.04**	**0.05**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
cDCC	**0.04**	**0.05**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
BEKK	**0.04**	**0.07**	**0.07**	**0.01**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OGARCH	**0.04**	**0.07**	**0.07**	**0.03**	**0.04**	**0.03**	**0.03**	**0.02**	**0.02**	**0.03**	**0.03**	**0.03**	**0.03**
GJR-CCC	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GJR-DCC	**0.04**	**0.07**	**0.01**	**0.01**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GJR-ADCC	**0.04**	**0.07**	**0.01**	**0.01**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GJR-cDCC	**0.04**	**0.07**	**0.01**	**0.01**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ABEKK	**0.04**	**0.07**	**0.01**	0.00	**0.04**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GJR-OGARCH	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**

Note: Bold shaded p-values denote models included in the confidence set for each of the problem dimensions reported in the first row at the 1% confidence level.