SFB 649 Discussion Paper 2009-007

Combination of multivariate volatility forecasts

Alessandra Amendola*
Giuseppe Storti*

*Department of Economics and Statistics,University of Salerno, Italy

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

SFB 649 ECONOMIC RISK BERLIN

Combination of multivariate volatility forecasts*

Alessandra Amendola [†] Giuseppe Storti [‡]

January 23, 2009

Abstract

This paper proposes a novel approach to the combination of conditional covariance matrix forecasts based on the use of the Generalized Method of Moments (GMM). It is shown how the procedure can be generalized to deal with large dimensional systems by means of a two-step strategy. The finite sample properties of the GMM estimator of the combination weights are investigated by Monte Carlo simulations. Finally, in order to give an appraisal of the economic implications of the combined volatility predictor, the results of an application to tactical asset allocation are presented.

Keywords: Multivariate GARCH, Forecast Combination, GMM, Portfolio Optimization.

JEL classification: C52, C53, C32, G11,G17.

1 Introduction

In banks and other financial institutions, the implementation of effective risk management strategies requires the creation and management of large dimensional portfolios. In theory multivariate GARCH (MGARCH) models offer a flexible tool for the estimation of portfolio volatility. In practice this is not the case if the dimension of the portfolio to be analyzed is even moderately (say > 10) large. The building of tractable multivariate models for the conditional volatility of high dimensional portfolios requires the imposition of severe constraints on the volatility dynamics. At the same time, data scarcity and computational

*Acknowledgements: This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk". The authors would like to thank participants to the CFE08 and IASC 2008 conferences for valuable comments and suggestions on a previous version of the paper.

[†] Department of Economics and Statistics,University of Salerno, Italy.
E-mail: alamendola@unisa.it

[‡] Department of Economics and Statistics, University of Salerno, Italy.
E-mail: storti@unisa.it

constraints limit the development of model selection techniques for non-nested multivariate volatility models. Hence, at the model building stage, constraints are often imposed on an *a priori* basis without following any formal statistical testing procedure. This situation leads to a potentially high degree of model uncertainty which can have a dramatic influence on the volatility predictions generated by different competing models. It is easy to recognize that this is a critical problem for risk managers and, in general, for any practitioner interested in the generation of accurate volatility forecasts. The problem of model uncertainty in multivariate conditional heteroskedastic models has already been addressed by Pesaran and Zaffaroni (2005). In order to reduce the risk deriving from inadvertently using a *wrong* MGARCH model, they discuss a procedure based on the use of Bayesian model averaging techniques. This paper proposes an alternative approach to dealing with model uncertainty in multivariate volatility predictions. Differently from Pesaran and Zaffaroni (2005), who focus on the combination of forecast probability distribution functions, our approach aims at combining point forecasts of conditional covariance matrices. The literature on the combination of conditional mean forecasts is quite mature dating back to the seminal paper by Bates and Granger (1969). Classical forecast combination techniques are based on the minimization of the Mean Squared Forecast Error (MSFE). So, in most cases, the combination weights associated to different competing models can be estimated by standard regression techniques. However, when combining volatility forecasts, loss functions such as the MSFE cannot be directly used since the conditional variance is not observed. So a proxy is needed. Common approaches rely on using squared returns, but these offer a noisy measure of volatility. An alternative solution is to use *realized volatility*, which is a much more accurate measure of volatility. Regarding the use of realized volatility, care is needed in the choice of the discretization interval. Too wide intervals result in inefficient estimates but, if the chosen integration interval is too narrow, micro-structure market frictions can distort the resulting measure of the unobserved volatility (Andersen et al. (2005)). Also, in some applications (e.g. macroeconomic applications) intra-daily (or, in general, *high frequency*) observations on the phenomenon of interest are not available and so realized volatility measures cannot be computed. Last, but not least, most of the literature on forecast combination typically deals with univariate time series while we are interested in the analysis of large dimensional multivariate processes.

To overcome these difficulties, in an univariate setting, Amendola and Storti (2008) have suggested a procedure for combining volatility forecasts which is based on the use of the Generalized Method of Moments (GMM) for the estimation of the combining weights. The moment conditions used to build the GMM criterion are based on theoretically founded restrictions on the stochastic structure of the standardized residuals.

Aim of this paper is to generalize this procedure to the combination of multivariate volatility forecasts. This task is not straightforward due to the dimensionality problems typically affecting multivariate conditional heteroskedastic models. In particular, it happens that the number of moment conditions to be imposed rapidly tends to explode with the model's dimension. This implies that the size of the problem becomes unmanageable even for relatively moderate values of the cross-sectional dimension. In order to overcome this problem, our approach is to disaggregate the full portfolio of assets into subsets of lower dimension. In practice, the estimation of combination weights is based on a two-step procedure. The first step is related to the combination of conditional covariance matrix forecasts for low-dimensional systems, namely bivariate systems. In this case the GMM estimation of the combination weights is performed following a direct generalization of the univariate procedure with the difference that, in a bivariate setting, we need to impose constraints not only on the autocorrelation functions of the raw and squared standardized residuals but also on their cross-correlations. In the following step, we apply a procedure which resembles, in the spirit, the *McGyver method* proposed by Engle (2007) for the estimation of high dimensional Dynamic Conditional Correlation (DCC) models. A complex computational problem is then disaggregated into a number of simpler low-dimension problems.

The structure of the paper is as follows. In section 2 the combined GMM volatility predictor is presented while the *disaggregate* estimation procedure for large dimensional systems is illustrated in section 3. The finite sample properties of the GMM estimator are investigated in in section 4 by means of a Monte Carlo simulation study while section 5 evaluates the economic relevance of the proposed procedure presenting the results of an application to portfolio optimization within a tactical asset allocation problem. The portfolio of assets we consider includes data on the whole set of 30 stocks used to compute the Dow Jones index. Some concluding remarks are given in the last section.

2 The combined GMM volatility estimator

In this section we introduce and discuss an approach to the combination of multivariate volatility forecasts generated by different, possibly non-nested, models. The Data Generating Process (DGP) is assumed to be given by

$$\begin{aligned} r_t &= x_t + u_t & (1) \\ u_t &= {H_t}^{1/2} z_t & (2) \end{aligned}$$

where r_t is a stationary and ergodic n-dimensional stochastic process; x_t is the conditional mean vector, which can potentially include lagged values of r_t as well as other regressors; z_t is a $(n \times 1)$ random vector with $E(z_t) = \mathbf{0}_{n,1}$ and $Var(z_t) = \mathbf{I}_{n,n}$, the order n identity matrix; ${H_t}^{1/2}$ is a $(n \times n)$ positive definite (p.d.) matrix such that

$${H_t}^{1/2}({H_t}^{1/2})' = var(r_t|I^{t-1}).$$

Assuming that a set of k candidate models for r_t is potentially available, let $\hat{x}_{ti}$, $i = 1, \ldots, k$, be the one step ahead predictor of r_t generated by the i^{th} model. The unconstrained *combined* predictor of the level of the r_t process can be defined as

$$\tilde{x}_t = \sum_{i=1}^{k} w_i^{(x)} \hat{x}_{ti}, \qquad (3)$$

with $\tilde{w}_i^{(x)} \in \Re$.

Similarly, let $\hat{H}_{t,i}$, for $i = 1, \ldots, k$, be the (1 step ahead) p.d. predicted covariance matrix generated by the i^{th} candidate model. The *combined conditional covariance matrix predictor* can be defined as

$$\tilde{H}_t = \sum_{i=1}^{k} w_i^{(h)} \hat{H}_{t,i}, \tag{4}$$

where $w_i^{(h)} \geq 0$ are the *combination weights* associated to each model. Also assume $\exists i : w_i^{(h)} > 0$, for $i = 1, \ldots, k$. The assumption of non-negative variance weights ($w_i^{(h)}$) is required in order to guarantee the *positive definiteness* of the combined volatility predictor.

Finally, it is important to remark that we choose not to impose the convexity constraint on the combining weights. The main advantage of adopting an unconstrained combination scheme is that it allows to yield an unbiased combined predictor even if one or more of the candidate predictors are biased. The standardized residuals from the combined volatility predictor are defined as

$$\tilde{z}_t = (\sum_{i=1}^{k} w_i^{(h)} \hat{H}_{ti})^{-1/2}(r_t - x_t). \tag{5}$$

At this stage the problem is how to estimate the optimal combination weights for the conditional mean and variance models. The approach we propose is based on the minimization of a GMM loss function implying appropriate (theoretically founded) restrictions on the moments of the standardized residuals $\tilde{z}_t$. Any specific choice of weights generates a different sequence of residuals characterized by different dynamical properties. The GMM estimator simply selects the vector of weights returning the sequence of residuals which most closely matches the theoretical restrictions imposed on $\tilde{z}_t$. Moments based estimators have already been used for estimating GARCH models parameters (Kristensen and Linton (2006); Storti (2006)). Other applications of the GMM approach in finance have been surveyed by Jagannathan et al. (2002). However the application of these techniques to the combination of multivariate volatility forecasts still deserves investigation.

Technically, the estimated combination weights ($\tilde{w}_i^{(x)}, \tilde{w}_i^{(h)}$) are chosen to solve the following minimization problem

$$\tilde{w} = \underset{w}{\operatorname{argmin}} \; m_T(w)' \hat{\Omega}_T^{-1} m_T(w) \tag{6}$$

where $\tilde{w} = (\tilde{w}^{(h)}, \tilde{w}^{(x)})'$ with $\tilde{w}^{(x)} = (\tilde{w}_1^{(x)}, \ldots, \tilde{w}_k^{(x)})$ and $\tilde{w}^{(h)} = \tilde{w}_1^{(h)}, \ldots, \tilde{w}_k^{(h)})$; $m_T(w) = \frac{1}{T}\sum_{t=1}^{T} \mu(w,t)$ and $\mu(w,t)$ is a $(N \times 1)$ vector of moment conditions; $\hat{\Omega}_T$ is a consistent p.d. estimator of

$$\Omega = \underset{T \to \infty}{lim} T E(m_T(w^*) m_T(w^*)')$$

with w^* being the solution to the moment conditions i.e. $E(\mu(w^*, t)) = 0$. Ω can be estimated by the heteroskedasticity and autocorrelation robust estimator proposed by Newey and West (1987). The weighting matrix $\hat{\Omega}_T$ plays an important role in GMM estimation.

Although its choice does not affect consistency, it can have dramatic effects on the efficiency of the GMM estimator (Newey and McFadden (1994)).
Each element of $\mu(w,t)$ specifies a restriction on the moments structure of the process $\tilde{z}_t$ in equation (5). In particular the vector $\mu(w,t)$ can be partitioned as follows

$$
\begin{aligned}
\mu_{i,t}^{(1)} &= \tilde{z}_{i,t} \qquad i=1,\dots,n.\\
\mu_{ij,t}^{(2)} &= \begin{cases} \tilde{z}_{i,t}^2 - 1 & \forall i=j;\\ \tilde{z}_{i,t}\tilde{z}_{j,t} & \forall i\neq j; i,j=1,\dots,n. \end{cases}\\
\mu_{ij,t}^{(3)} &= \tilde{z}_{i,t}\tilde{z}_{j,t-h} \qquad h=1,\dots,g.\\
\mu_{ij,t}^{(4)} &= \tilde{z}_{i,t}^2\tilde{z}_{j,t-h}^2 - 1 \quad h=1,\dots,g.
\end{aligned}
$$

The rationale behind the choice of the above reported moment conditions is to constrain the standardized residuals $\tilde{z}_t$, implied by a given set of combination weights, to be as close as possible to a sequence of i.i.d. random vectors with zero expectation and identity covariance matrix. The first set of conditions ($\mu_{i,t}^{(1)}$) restricts the standardized residual to have zero expectation, $E(\tilde{z}_t) = \mathbf{0}_{n,1}$. The conditions on the covariance matrix, $E(\tilde{z}_t\tilde{z}_t') = \mathbf{I}_{n,n}$, are met through $\mu_{ij,t}^{(2)}$. The other two set of conditions $\mu_{ij,t}^{(3)}$ and $\mu_{ij,t}^{(4)}$ respectively imply that $E(\tilde{z}_t\tilde{z}_{t-h}') = \mathbf{0}_{n,n}$ and $E[\tilde{z}_t^{\cdot 2}(\tilde{z}_{t-h}^{\cdot 2})'] = \mathbf{1}_{n,n}$, where $\tilde{z}_t^{\cdot 2}$ is the vector obtained by squaring each element in $\tilde{z}_t$. The last set of conditions simply implies that $\tilde{z}_{i,t}^2$ and $\tilde{z}_{j,t-h}^2$ are uncorrelated $\forall i,j=1,\dots,n$.

One difficulty with the direct application of this approach to large dimensional systems is that the number of moment conditions to be imposed rapidly increases with the model's cross-sectional dimension n . This relationship is graphically represented in Figure 1 for the case $g=1$. So an appropriate strategy for reducing the problem to a tractable dimension is needed. This point is addressed in the following section.

3 Disaggregate estimation of the combination weights

This section illustrates a two-step procedure which allows to apply our GMM procedure for the combination of volatility forecasts to the modeling of large dimensional portfolios. In the spirit of the *"MacGyver"* method, proposed by Engle (2007) for the estimation of high dimensional DCC (Engle (2002)) models, we extract all the possible bivariate systems $(r_{i,t}, r_{j,t})'$ from the n-variate returns process ($\forall i,j$). Then we use the GMM estimator in (6) to generate a set of consistent estimates of the weights vector ($\tilde{w}_{i,m}^{(h)}; \tilde{w}_{i,m}^{(x)}$) for each bivariate subsystem (m=$1,\dots,n(n-1)/2$; $i=1,\dots,k$). For the i-th candidate model, the resulting set of estimates is

$$W(i,M;2) = \left(\tilde{w}_{i,1}^{(M)},\dots,\tilde{w}_{i,P}^{(M)}\right) \qquad P=n(n-1)/2,\ i=1,\dots,k;\ M=h,x$$

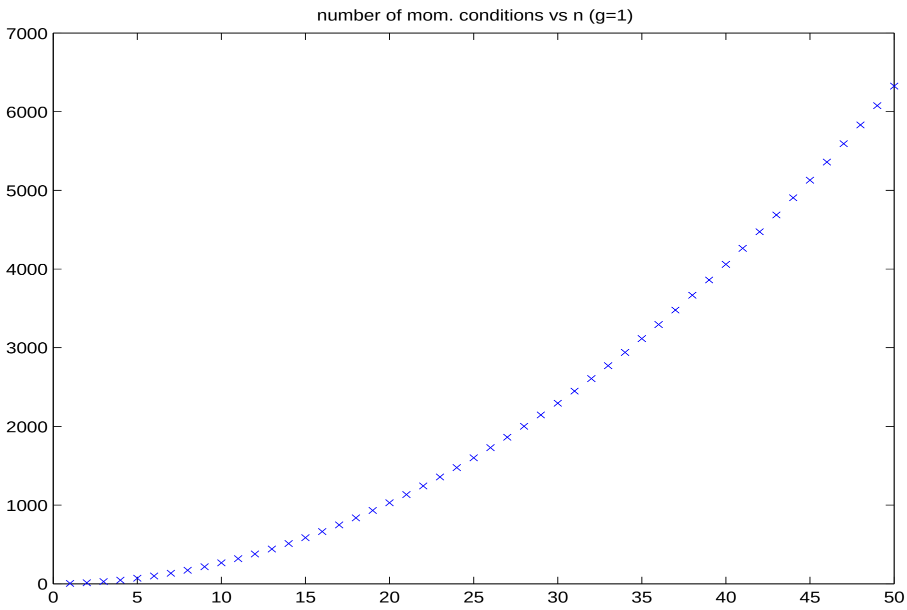

Figure 1: Number of moment conditions N vs n ($g = 1$).

The final estimates of the $w_i^{(M)}$ are computed by applying a *blending* function $B(.)$ to the set $W(i, M; 2)$

$$B(W(i, M; 2)) = \tilde{w}_i \qquad i = 1, \ldots, k$$

The *blending* function $B(.)$ can be any function satisfying the following consistency requirement

$$B(w_i^{(M)}, \ldots, w_i^{(M)}) = w_i \qquad i = 1, \ldots, k; M = h, x.$$

The asymptotic properties of this estimator are not analytically known although Engle (2007) investigates its finite sample properties by means of Monte Carlo simulations with a cross-sectional dimension n varying from 3 to 50. The simulation results indicate that the most accurate results are obtained when the median is used as blending function. For this reason we also use the median of the bivariate estimates as *blending* function in our estimations.

Using the disaggregate estimation procedure for dealing with large datasets gives a relevant practical advantage. It allows to drastically reduce the number of moment conditions to be simultaneously handled (which is equal to the corresponding number for a bivariate system) giving a feasible solution to the estimation problem for large values of n. This is evident from Table 3 which reports the number of moment conditions reached in the *disaggregate procedure* as a function of the maximum lag value g. Another advantage of this

g	1	5	10	15	20
n_m	13	45	85	125	165

Table 1: Number of moment conditions for the bivariate system vs. g

approach is that, in very large dimensional systems, it is not necessary to analyze the whole set of bivariate systems, whose number can be overwhelming, but it is in theory possible to perform the estimation only on a subset of $P_0 < P$ bivariate systems. This implies that, in situations in which the structure of the portfolio is continuosly changing over time, the weights do not need to be re-estimated each time a new asset is added or excluded from the portfolio. There is however no guidance on how to optimally select the subset of assets to be used for estimation. Also, although the disaggregate estimation procedure by its own definition returns consistent estimates, its theoretical efficiency properties are not analytically known. Finally, the empirical distribution of the $\tilde{w}_i^{(m)}$ can provide useful information for detecting misspecifications. For example, a very disperse distribution can provide hints in favor of the presence of heterogeneity in the combining weights.

4 A simulation experiment

In order to investigate the finite sample properties of the GMM estimator of the weights $\tilde{w}_j$ ($j = 1, \ldots, 2k$) we perform a Monte Carlo simulation study considering two different settings. In the first case the DGP is assumed to be given by a bivariate system ($n = 2$) while, in the second, it is assumed to be a system of dimension $n = 20$. In both cases, the number of candidate models to be combined is set to k=2. Namely, we use a scalar DCC and a scalar VEC (Bollerslev et al. (1988)) model. The conditional mean is assumed to be equal to zero. The updating equation for the conditional covariance matrix H_t implied by the chosen DCC model is defined by the following equations

$$\begin{aligned}
H_{t,DCC} &= D_t R_t D_t \\
D_t &= \mathrm{diag}(H_t^*) \qquad H_{ii,t}^* = \sqrt{V_{ii,t}} \\
V_{ii,t} &= a_{0,i} + a_{1,i} r_{i,t-1}^2 + b_{1,i} V_{ii,t-1} \\
R_t &= (\mathrm{diag}(Q_t))^{-1} Q_t (\mathrm{diag}(Q_t))^{-1} \\
Q_t &= R(1 - 0.02 - 0.96) + 0.02(\epsilon_{t-1}\epsilon_{t-1}^{'}) + 0.96 Q_{t-1}.
\end{aligned}$$

with i=$1, \ldots, n$, $\epsilon_t = D_t^{-1} r_t$, $R = \mathrm{corr}(\epsilon_t)$. For the VEC model, the updating equation for H_t is

$$\mathrm{vech}(H_{t,VEC}) = \underline{c} + 0.03\mathrm{vech}(r_{t-1} r_{t-1}^{'}) + 0.95\mathrm{vech}(H_{t-1,VEC}).$$

Since we have set $x_t = 0$, $\forall t$, we only need to specify the conditional variance weight associated to each candidate model. The DGP is then defined by the following two equations

$$r_t = H_t^{1/2} z_t \tag{7}$$

$$H_t = w_{DCC}^{(h)} H_{t,DCC} + w_{VEC}^{(h)} H_{t,VEC} \tag{8}$$

with $w_{DCC}^{(h)} = 0.65$, $w_{VEC}^{(h)} = 0.35$ and $z_t \underset{iid}{\sim} MVN(0_{n,1}, I_{n,n})$. Direct GMM estimation of the weights is feasible only for $n = 2$ while, for $n = 20$, we need to resort to the disaggregate

estimation procedure described in the previous section. First, we estimate the weights for each of the possible 190 bivariate subsystems. Second, the final estimates of $\{w_{DCC}^{(h)}, w_{VEC}^{(h)}\}$ are computed taking the medians of the empirical distributions of the resulting bivariate estimates.
The simulation study has been repeated for four different sample sizes, namely T={500, 1000, 2000, 5000}. For each sample size, in the bivariate case, we have generated 500 independent Monte Carlo replicates while, due to computational constraints, in the case of $n = 20$ only a single series is generated for each value of T.

The maximum lag used to build the GMM moment conditions has ben set to $g = 1$ and the optimal weighting matrix Ω is estimated by the Newey-West estimator (Newey and West (1987)). One complication arising with the Newey-West estimator is that the estimated asymptotic covariance comes to depend on the parameter vector w. To overcome this difficulty, following common practice in the GMM literature, we adopt a two stage estimation procedure. First, we set $\Omega = I_{N,N}$ to generate an initial consistent estimate of the parameter vector $\tilde{w}^{\dagger}$. Second, we use $\tilde{w}^{\dagger}$ to generate a consistent estimator of Ω, $\hat{\Omega}(\tilde{w}^{\dagger})$ which is plugged into (6). A more efficient estimator of w is then obtained from the maximization of the resulting loss function.
For the bivariate case, the simulated distributions of the estimated combination weights are summarized in Figure 2 by means of box-plots. It can be easily observed how the bias component is negligible for any value of T while the variability of the estimates is rapidly decreasing as T is increasing. Similar considerations hold for the high dimensional

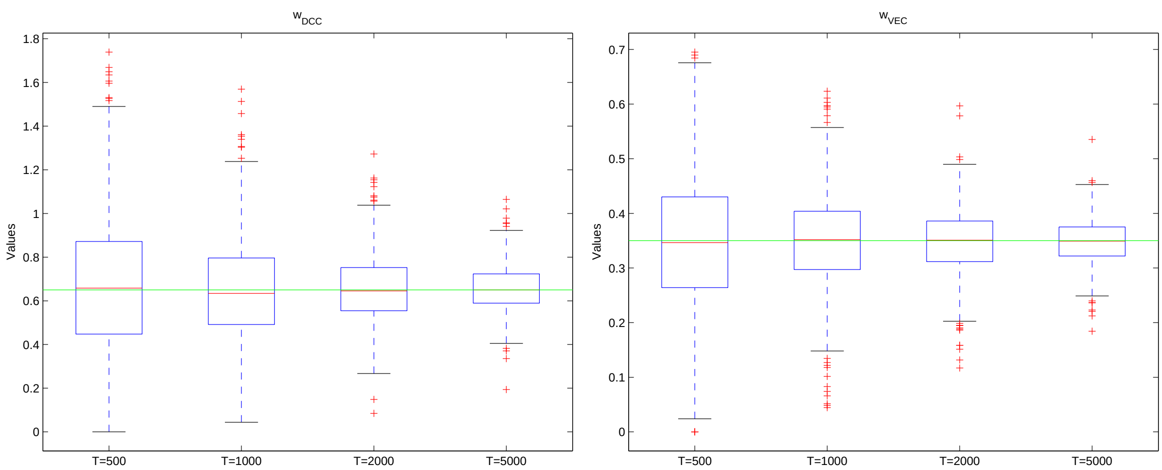

Figure 2: Simulation results for bivariate systems: : $\tilde{w}_{DCC}$, left, and $\tilde{w}_{VEC}$, right. The (green) horizontal line indicates the true parameter value.

case. Figure 3 reports the box-plots of the estimated weights computed from each of the 190 feasible bivariate subsystems. There is some bias for $T = 500$ but this is rapidly disappearing for higher sample sizes. Also the distribution of bivariate estimates tends to be characterized by lower variability as T increases. It is important to note that, although our exercise provides evidence in favor of the use of the disaggregate estimation procedure, the empirical distributions in Figure 4 cannot be interpreted as estimates of the sampling distribution of the disaggregate estimator. This is due to the fact that they are referred to a single simulated series.

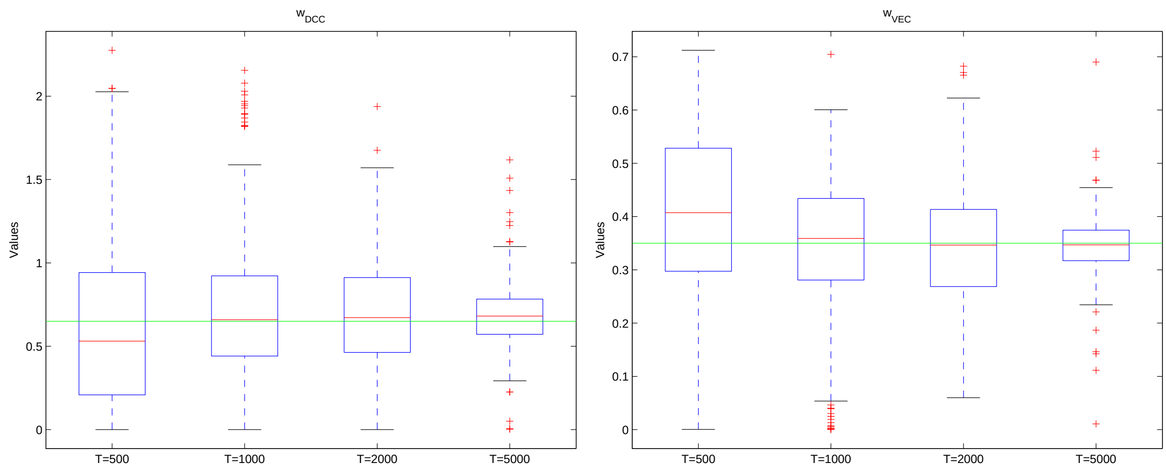

Figure 3: Simulation results for high-dimensional systems ($n = 20$): $\tilde{w}_{DCC}$, left, and $\tilde{w}_{VEC}$, right. The (green) horizontal line indicates the true parameter value

5 Empirical evidence on financial data: an application to tactical asset allocation

In order to assess the effectiveness of the proposed procedure in real financial applications, in this section we present the results of an application to a portfolio optimization problem. Namely, we consider a mean-variance framework in which we fix a target expected return and try to minimize the portfolio volatility (see e.g. Fleming et al. (2001)). The portfolio we consider is composed of a basket of risky assets, the whole set of stocks included in the Dow Jones stock market index, and a riskless asset, a 3 months constant maturity US Treasury bill.

We consider daily data ranging from 11.01.1999 to 11.08.2008 for a total of 2501 datapoints. For stocks, returns were calculated as the first difference of log-transformed (adjusted) daily closing prices while returns on the risk free asset were measured in terms of the interest rate on the 3 months US Treasury bill, adjusting for weekends and holidays. All the data were downloaded from Datastream. Again, as candidate models, we select a DCC and a VEC model. The conditional mean series x_t is assumed to be constant. The dataset is divided into three parts, observations from 1 to 1000 are used to generate initial estimates of the parameters of the two candidate models. Conditional on these estimates, we generate volatility forecasts for observations from 1001 to 2000. These predictions are

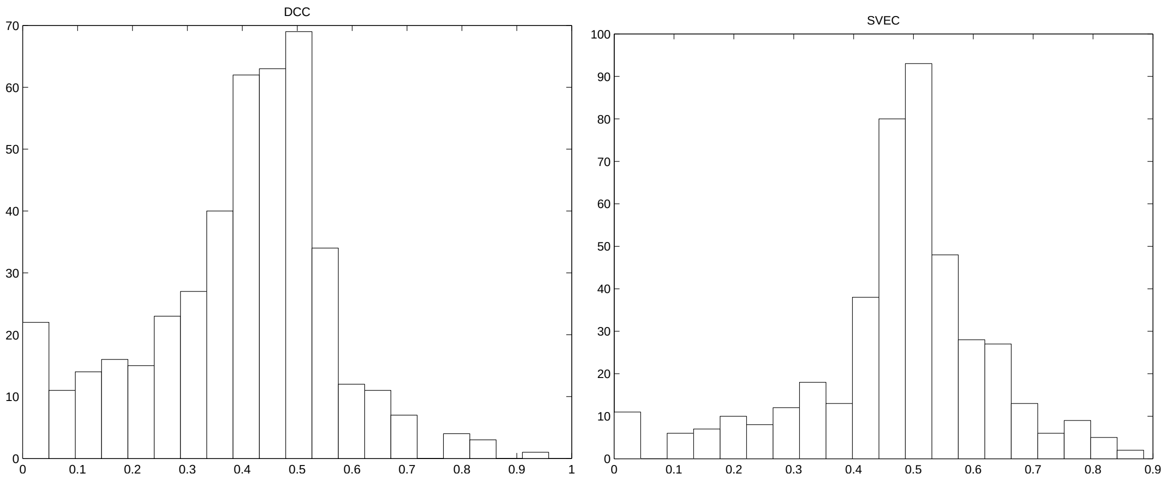

Figure 4: Estimated weights distribution over 435 bivariate subsystems

then used to estimate the combination weights and to calculate the optimal combined volatility predictor. Finally, observations from 2001-2500 are used for out-of-sample forecast evaluation.

Namely, the first step is to fit the candidate models to data points included in period 1 and use the estimated parameters to generate 1-day-ahead volatility predictions for period 2. The estimated models are

$$Q_t = R(1 - 0.0060 - 0.6714) + \underset{(0.0023)}{0.0060}(\epsilon_{t-1}\epsilon_{t-1}^{'}) + \underset{(0.2602)}{0.6714}Q_{t-1}$$

$$H_t = S(1 - 0.0080 - 0.9563) + \underset{(0.0008)}{0.0080}(u_{t-1}u_{t-1}^{'}) + \underset{(0.0050)}{0.9563}H_{t-1}$$

with $S = var(r_t)$. Volatility predictions generated from these models are used to estimate the optimal combination weights ($\tilde{w}$).

Relying on the estimated weights $\tilde{w}$ we generate volatility predictions for period 3. These predictions are then used for determining the optimal portfolio allocation over the same period. Finally, the results of step 3 are compared with those obtained by separately estimating the 2 candidate models using data from period 2. In this way both the combined predictor and the individual models used for comparison are based on "concurrent" information sets.

The distributions of the estimated combination weights $\tilde{w}_i$ over 435 bivariate subsystems are reported in Figure 5. The medians of these distributions are equal to 0.4192, for the DCC, and to 0.4921 for the VEC model. The vector of optimal portfolio weights at time t ($\hat{\omega}_t$) is obtained as a solution of the constrained optimization problem

$$\underset{\omega_t}{\text{argmin}} \quad \omega_t^{'} H_t \omega_t$$

subject to

$$\omega_t^{'}\mu + (1 - \omega_t^{'}u)r_{f,t} = \mu_p$$

where $u = \mathbf{1}_{n,1}$, $\mu = E(r_{t+1})$, μ_p is the expected target rate of return and $r_{f,t}$ is the daily return rate on the risk free asset. The solution is analytically found as

$$\hat{\omega}_t = \frac{(\mu_p - r_{f,t})H_t^{-1}(\mu_p - r_{f,t}u)}{(\mu - r_{f,t}u)' H_t^{-1}(\mu - r_{f,t}u)} \tag{9}$$

So, at each time point, the portfolio weights at time $(t+1)$ are recalculated as a function of the current prediction of the future conditional variance matrix H_{t+1}. In our exercise we allow for short selling (and so for negative portfolio weights) while we do not consider the effect of transaction costs.

Using equation (9) we have computed the optimal portfolios based on the combined predictor and the candidate models estimated using data from period 2. We have then compared their performances in terms of mean and variance of the implied portfolio returns $r_{p,t}$, autocorrelation function of standardized portfolio returns, final wealth (W) $\times$ unit investment, expected utility (U), Sharpe ratio (SR).

The expected utility is calculated as in Fleming et al. (2001)

$$U(\gamma) = \frac{1}{T}\sum_{t=0}^{T-1}\left[r_{p,t+1} - \frac{\gamma}{2(1+\gamma)}r_{p,t+1}^2\right]$$

In the above formula the constant γ can be interpreted as a measure of the investor's relative risk aversion. The results are summarized in Table 2. We consider two different values of the annual target return μ_p, 0.10 and 0.20. In both cases, the VEC model returns the portfolio with the minimum variance. The portfolio implied by the combined predictor is characterized by a slightly higher variance but also by the highest average return, final end of the period wealth and Sharpe ratio. The value of expected utility measure gives a summary measure of the overall performance of the investment strategies associated to each competing model. We compute the expected utility for two different values of the risk aversion parameter γ, 1 and 10. In both cases the highest expected utility is that returned by the portfolio associated to the investment strategy implied by the combined predictor.

As a further benchmark for evaluating the perfomances of the different approaches, for each strategy, we consider the autocorrelation functions of the implied standardized portfolio residuals

$$z_{tj(p)} = r_{t(p)}/\hat{h}_{tj(p)} \qquad j = 0, 1, \ldots, k$$

where, letting $\hat{H}_{t0} = \tilde{H}_t$ in order to compact notation, $\xi_{tj}' = \begin{pmatrix}\hat{\omega}_{tj}' & 1 - \hat{\omega}_{tj}u\end{pmatrix}$ is the vector of portfolio weights implied by model j and $\hat{h}_{tj(p)}^2 = \xi_{tj}'\hat{H}_{tj}\xi_{tj}$ is the portfolio variance implied by the $j-th$ candidate model. In Figure 5 we report the p-values of the Ljung-Box Q-test performed on the sample autocorrelation functions of $z_{tj(p)}$ and $z_{tj(p)}^2$, taking into account lags from 1 to 10. For $z_{tj(p)}$, both the DCC and VEC portfolios show significant serial correlation at lag 1. For the DCC model, there is also evidence of autocorrelation at lag 10. The portfolio generated by the combined predictor is not affected by this problem. When we move to consider the autocorrelation function of $z_{tj(p)}^2$, we note that the DCC portfolio's squared residuals are characterized by a significant autocorrelation

	$mu_p = 0.10^+$		
	comb.	DCC	VEC
$var(r_{p,t})^*$	0.0559	0.0570	0.0545
$mean(r_{p,t})^+$	0.0582	0.0216	0.0362
W	1.1209	1.0424	1.0730
$U(1)^*$	2.2973	0.8442	1.4229
$U(10)^*$	2.2857	0.8325	1.4117
SR	0.0731	0.0115	0.0366
	$mu_p = 0.20^+$		
	comb.	DCC	VEC
$var(r_{p,t})*$	0.2612	0.2647	0.2527
$mean(r_{p,t})^+$	0.1123	0.0324	0.0649
W	1.2414	1.0594	1.1304
$U(1)^*$	4.6633	1.3003	2.6301
$U(10)^*$	4.3370	1.1652	2.4622
SR	0.0758	0.0137	0.0397

Table 2: Summary statistics for alternative portfolio allocations based on the combined volatility predictor, DCC and VEC models. Legend: ($^+$) annualized value; (*) $\times\ 10^4$
.

pattern providing evidence that the DCC model is not able to characterize the volatility dynamics of the portfolio returns. Again, this problem disappears when the conditional covariance matrix estimates generated by the combined predictor are used to determine the optimal portfolio.

6 Concluding Remarks

In multivariate modelling of conditional volatility for large dimensional portfolios, model identification is not an easy task due to data scarcity and computational constraints. In this framework, combining volatility forecasts from different models offers a simple and practical solution for dealing with model uncertainty avoiding the risks related to having to select a single candidate model.

Also, the two-step GMM approach to the estimation of the combination weights which is discussed in this paper allows to deal with the prediction of volatility matrices in high dimensional systems, overcoming the *curse of dimensionality* problem typically arising in MGARCH models.

The results of our Monte Carlo simulation study provide encouraging evidence on the finite sample properties of the proposed procedure in terms of both bias and variance. Finally, the results of an application to a portfolio optimization problem suggest that our GMM approach to combining volatility forecasts can be effectively applied in routine risk

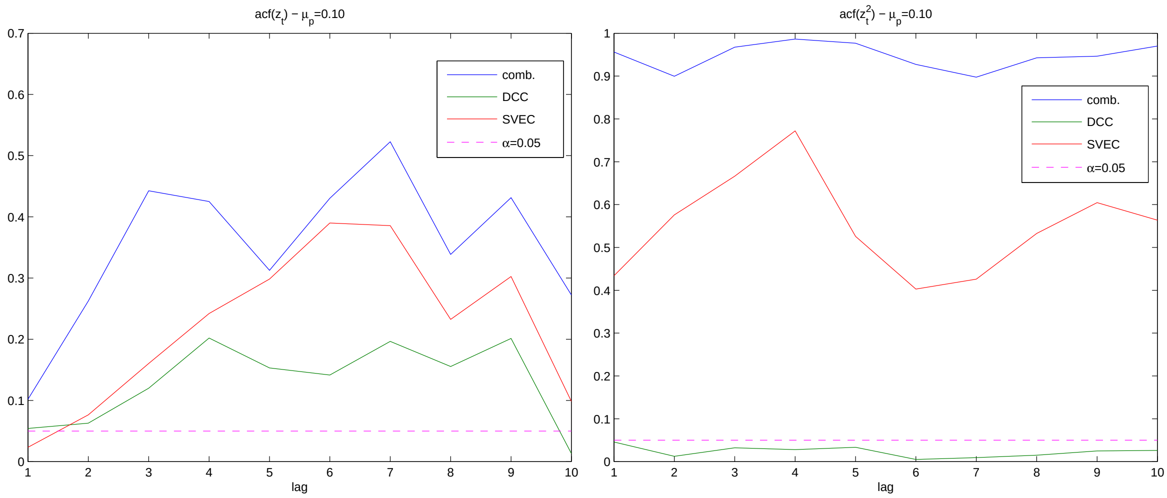

Figure 5: p-values vs. lag of the Ljung-Box Q-statistic for the standardized portfolio residuals from different models ($z_{p,t}$, left) and their squares ($z^2_{p,t}$, right) ($\mu_p = 0.10$). Similar results are obtained for $\mu_p = 0.20$.

management applications, allowing to improve over the performance of single (possibly misspecified) volatility models.

References

Amendola, A., Storti G. (2008), A GMM procedure for combining volatility forecasts. *Computational Statistics & Data Analysis*, 52(6), 3047-3060.

Andersen, A. T., Bollerslev, F., Meddahi, N. (2005) Correcting the errors: a note on volatility forecast evaluation based on high frequency data and realized volatilities, *Econometrica*, 73, 279-296.

Bates, J. M, Granger, C. W. J. (1969) The combination of forecasts, *Operational Research Quarterly*, 20, 319-325.

Bollerslev, T.P., Wooldridge, J.M., Engle, R. (1988) A Capital Asset Pricing Model with Time Varying Covariances, *Journal of Political Economy* 96 , 116-131.

Engle, R. (2002), Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models,*Journal of Business & Economic Statistics*, 20, 339350.

Engle, R. (2007) High Dimension Dynamic Correlations, *Working Paper*.

Fleming, J., Kirby, C., Ostdiek, B. (2001) The economic value of volatility timing, *Journal of Finance*, 56, 329-352.

Jagannathan, R., Skoulakis, G., Wang, Z. (2002) Generalized Method of Moments: Applications in finance, *Journal of Business and Economic Statistics*, 20, 470-81.

Kristensen D., Linton O. (2006) A Closed-form Estimator for the GARCH(1,1) model, *Econometric Theory*, 22, 323-327.

Newey, W. K., McFadden, D. (1994) Large sample estimation and hypothesis testing, in *Handbook of Econometrics*, Volume IV, 2113-2245, Elsevier Science.

Newey, W. K., West, K. D. (1987) A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55, 3, 703-708.

Pesaran, H. M., Zaffaroni, P. (2005) Model averaging and Value-at-Risk based evaluation of large multi-asset volatility models for risk management, *CEPR Discussion Paper* No. 5279 .

Storti, G. (2006) Minimum distance estimation of GARCH(1,1) models, *Computational Statistics & Data Analysis*, 51, 1803-1821.

SFB 649 Discussion Paper Series 2009

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Implied Market Price of Weather Risk" by Wolfgang Härdle and Brenda López Cabrera, January 2009.
002 "On the Systemic Nature of Weather Risk" by Guenther Filler, Martin Odening, Ostap Okhrin and Wei Xu, January 2009.
003 "Localized Realized Volatility Modelling" by Ying Chen, Wolfgang Karl Härdle and Uta Pigorsch, January 2009.
004 "New recipes for estimating default intensities" by Alexander Baranovski, Carsten von Lieres and André Wilch, January 2009.
005 "Panel Cointegration Testing in the Presence of a Time Trend" by Bernd Droge and Deniz Dilan Karaman Örsal, January 2009.
006 "Regulatory Risk under Optimal Incentive Regulation" by Roland Strausz, January 2009.
007 "Combination of multivariate volatility forecasts" by Alessandra Amendola and Giuseppe Storti, January 2009.

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".