An Out-of-sample Analysis of Mean-Variance Portfolios with Orthogonal GARCH Factors

Alessandro Cardinali®

University of Bristol

ABSTRACT

In this paper a comparative study is conducted to evaluate the out-of-sample performance of mean-variance portfolios when three different variance models are considered. We use the common framework of orthogonal factors to specify the conditional covariance matrix structure. A key advantage of this approach is that the estimated factors can be modeled as univariate GARCH processes so that we can consider models for which multivariate extensions are not available. We, therefore, compared the Integrated GARCH (IGARCH) with the Exponential GARCH (EGARCH) and Fractionally Integrated Exponential GARCH (FIEGARCH) factor models on the basis of statistical diagnostics, and found the EGARCH model superior when fitted with heavy tailed distributions. We also evaluated out-of sample portfolio performances in terms of efficient frontiers, prediction intervals and turnover, and concluded that the EGARCH and FIEGARCH models provide comparable outcomes which are overall superior to the IGARCH performance. Looking jointly at statistical and economic criterions we conclude that fitting a FIEGARCH model with heavy tailed distributions can generally improve out-of-sample portfolio performances.

Key words: *Mean-Variance Portfolios, GARCH Processes, Forecasting, Turnover*
JEL Classifications: C52, G11, G17

1. INTRODUCTION

Portfolio selection is the process by which an investor decides how to allocate the wealth among a universe of financial assets. Asset portfolios can be classified in two categories according to the purpose of their selection and, consequently, the different type of assets included. We can distinguish between:

- Asset allocation portfolios: typically include market indices as indicators of financial markets behavior. Their aim is to determine the optimal investment sharing among different markets;

- Equity portfolios: usually include a large number of equities from a certain market. Their aim is to select the optimal combination of single assets within a given market.

The two selection strategies are complementary; the first one establishes investment sharing between several markets, while the second step selects equities within each market. The statistical approach to this selection problem treats asset prices and returns as random variables and the investor perception of preferences is usually formalized defining a utility scale. Most of the approaches to portfolio selection depend on the expected utility of the final

® Alessandro Cardinali, Research Associate in Statistics, University of Bristol, UK. Email: A.Cardinali@bristol.ac.uk.

wealth, which is maximized to determine an optimal set of weights. This expectation is typically taken over the multivariate probability distribution of the asset returns. Maximization of the expected utility, therefore, leads to a criterion that depends on the parameters of the underlying probability distribution of returns. If multivariate normally distributed returns are assumed, these parameters are the vector of expected returns and the variance-covariance (VC) matrix. Even in the case of non-Gaussian returns, the VC matrix is important in a portfolio selection process and needs to be estimated or forecasted. Since temporal dependence is often observed in return's distribution, a conditional approach can be appropriate to estimate and forecast the mean vector and the VC matrix as inputs in the portfolio selection process. For this purpose, the use of GARCH predictions is very popular in the quantitative finance literature. As a recent example, Hawkes and Date (2007) compare several GARCH forecasts using statistical measures. In contrast to the large body of literature on statistical comparison of GARCH-type predictions, not many studies consider comparisons that directly evaluate portfolio performances. Although it is well known that portfolio selection is more sensitive to estimation error in expected returns (Chopra and Ziemba, 1993), portfolio weights and subsequent portfolio performance are also sensitive to the VC matrix estimation. In practice, the most important portfolio selection application is to ensure the optimal portfolio prediction over a certain horizon.

The aim of this paper is to assess the sensitivity of portfolio out-of-sample performances using measures from statistics and quantitative economics. In particular, we will investigate the relative merits of some VC matrix time series based estimation and forecasting approaches. In a number of studies (see Adcock, 2004) the use of factor models has been associated with the skewed Normal distribution to model asset returns. Some investigation has been conducted using Orthogonal GARCH factors to analyze the asset covariance matrix (Bystrom, 2004), but these analyses have not included evaluation of portfolio performances. Our paper links various aspects of the aforementioned literature. We conduct an empirical investigation of portfolios by means of factor models. Moreover, we explicitly consider skewed GARCH evolution for orthogonal factors and evaluate the relative merits of incorporating heavy tails and long memory.

The paper is structured as follow. Section 2 reviews the conditional mean-variance portfolio selection, deriving the mean-variance efficient frontier and introducing the relevant notation. Section 3 reviews the definition of Orthogonal GARCH factors for conditional VC matrix estimation and forecasting. This description also includes a review of univariate GARCH specifications that will be compared in the following analyses. Section 4 contains an empirical analysis which includes estimation of orthogonal variance factors, along with the fitting of three competing GARCH model specifications and their statistical diagnostics. Section 5 compares the out-of-sample portfolio performances obtained by considering different factor models. We mainly look at the empirical size of prediction intervals and at portfolio turnover as measure of risk management cost. We show these results according to an increasing scale of risk appetite. Section 6 concludes and outlines the direction for future research.

2. MEAN-VARIANCE PORTFOLIOS

The classical portfolio selection method – also known as mean-variance approach (Markowitz, 1952) – has been widely used by the financial community and is, in principle, not strictly related to the maximized expected utility. However, under certain conditions and assuming a quadratic utility function, mean-variance selection can be viewed as optimal in terms of investor's preferences. A simple way to summarize asset return's distribution (both analytically and graphically) is to represent portfolios in the mean-variance plan. The portfolio mean is thus assumed to be a measure of expected gain while the portfolio variance is considered as a measure of risk. In its original version this is a static model based on the following assumptions:

- The distribution of asset returns (and linear factors) is time invariant so that no time dependence is involved in moments estimation;

- The selection problem is viewed as one-period optimal, so at the end of each period the selection has to be updated.

Typically, the return's distribution is not time invariant and out-of-sample evaluations require a conditional approach, so that the portfolio selection will be performed using all information available up to a certain time. However, even in this more general setup, the selection problem is, in principle, viewed as one-period optimal. In practice the costs of renegotiating a new portfolio need to be compared with the expected increase of wealth, so the weights are not always conveniently updated. In other words, not always statistical and economical optimality coincide. This will lead us to consider both aspects in our comparative study. As shown in previous studies (Best and Grauer, 1991), the portfolio selection can be directly formulated as a Quadratic Programming Problem (QPP). For $\tau > 0$, setting $u = t + \tau$, when a single budget constraint (imposing weights summing-up to unity) is considered, a closed form solution is available and is given by

$$\begin{aligned}\mathbf{w}_{t,u;q} &= \arg\max_{\mathbf{w}} \{q\boldsymbol{\mu}'_{t,u}\mathbf{w} - ½\mathbf{w}'\,\boldsymbol{\Sigma}_{t,u}\mathbf{w}|\mathbf{1}'\,\mathbf{w} = 1\} \\ &= \boldsymbol{\Sigma}^{-1}_{t,u}\mathbf{1}/a_{t,u} + q[\boldsymbol{\Sigma}^{-1}_{t,u}\,(\boldsymbol{\mu}_{t,u} - \mathbf{1}b_{t,u}/a_{t,u})],\end{aligned} \tag{2.1}$$

where $\mathbf{x}_u$ is a vector of p asset returns at time $u > t$, $\boldsymbol{\mu}_{t,u}=\mathbb{E}(\mathbf{x}_u|\Omega_t)$, $\boldsymbol{\Sigma}_{t,u}= Var(\mathbf{x}_u|\Omega_t)$, $a_{t,u}= \mathbf{1}'\boldsymbol{\Sigma}^{-1}_{t,u}\mathbf{1}$ and $b_{t,u} = \mathbf{1}'\boldsymbol{\Sigma}^{-1}_{t,u}\boldsymbol{\mu}_{t,u}$. The constant q represents the *risk appetite* parameter, an index upon which the optimal portfolio solution will depend. Allowing the risk appetite q to vary, we obtain the set of optimal portfolios previously referred to as *efficient frontier*. Along the frontier, the choice of an optimal portfolio depends on the individual preferences, being relevant in terms of risk appetite/aversion, and represented by the value of the parameter q. The efficient portfolios mean and variances can therefore be easily derived. Using vector notation we have

$$m_{t,u;q} = \boldsymbol{\mu}'_{t,u}\mathbf{w}_{t,u;q} = \beta_0 + \beta_1 q, \tag{2.2}$$

for some constants β_0, β_1. The interested reader can refer to Best and Grauer (1991) for further details concerning this section. Equation (2.2) shows that the efficient portfolios mean is a linear function of the risk appetite coefficient. On the other hand

$$v_{t,u;q} = \mathbf{w}'_{t,u;q}\,\boldsymbol{\Sigma}_{t,u}\mathbf{w}_{t,u;q} = \gamma_0 + \gamma_1 q^2, \tag{2.3}$$

for some constants γ_0, γ_1, so that the efficient portfolio variance is a quadratic function of the risk appetite parameter. Solving the equation (2.2) for q and substituting in (2.3), we have the analytic expression for the conditional efficient frontier that is

$$(m_{t,u;q} - \beta_0) = \beta_1(v_{t,u;q} - \gamma_0). \tag{2.4}$$

This is a parabola in mean-variance plan or a hyperbola in mean-standard deviation space.

3. ORTHOGONAL FACTORS VARIANCE (OFV) MODELS

This section reviews the definition of Orthogonal GARCH factors for conditional VC matrix estimation and forecasting. This description also includes a review of univariate GARCH specifications that will be compared in the following analyses. The Orthogonal GARCH model was first proposed in Alexander (2001), and is based on Principal Component Analysis (PCA). A spectral decomposition is conducted on the estimated covariance matrix returning linear combinations of log-returns (the so called risk factors) being uncorrelated by construction and having their variances represented by their eigenvalues. Consequently, the usual advantages of PCA can be used to facilitate the covariance matrix estimation. Furthermore, the estimated matrix will be positive definite by construction, if all factors are considered. Most importantly, predictions can be obtained by using univariate models. To introduce the relevant notation, we can consider the problem dealt with render diagonal the full square and symmetric matrix $\mathbf{\Sigma}_t = Var(\mathbf{x}_t|\Omega_t)$. The corresponding eigensystem can be written as

$$\mathbf{\Sigma}_t\mathbf{P}_t = \mathbf{P}_t\mathbf{\Lambda}_t\,, \tag{3.5}$$

where $\mathbf{P}_t$ is the normalized eigenvector matrix and $\mathbf{\Lambda}_t$ is the eigenvalues matrix. By construction, the first is an orthogonal and positive definite matrix and the second is diagonal positive semi-definite. The covariance matrix spectral decomposition is therefore given by $\mathbf{\Sigma}_t = \mathbf{P}_t\mathbf{\Lambda}_t\mathbf{P}_t'$. Omitting the classical details in PCA derivation, we will indicate the principal components as the vector $\mathbf{y}_t = \mathbf{P}_t'\mathbf{x}_t$., the conditional mean vector as $\mathbb{E}(\mathbf{y}_t|\Omega_t) = \mathbf{g}_t$ and the conditional covariance matrix as $Var(\mathbf{y}_t|\Omega_t) = \mathbf{\Lambda}_t$. The eigenvalues will be in decreasing order and will represent the proportion of variance explained by the orthogonal factors. Orthogonality allows the use of univariate models to forecast the factors conditional mean vector as $(\mathbf{y}_u|\Omega_t) = \mathbf{g}_{t,u}$ and the factors covariance matrix as $Var(\mathbf{y}_u|\Omega_t) = \mathbf{\Lambda}_{t,u}$. It follows that we can define the assets conditional mean vector as $\boldsymbol{\mu}_{t,u} = \mathbf{P}_t\mathbf{g}_{t,u}$ and conditional covariance matrix as $\mathbf{\Sigma}_{t,u} = \mathbf{P}_t\mathbf{\Lambda}_{t,u}\mathbf{P}_t'$.

3.1. GARCH Factors

We assume the p orthogonal factors to be modeled as $y_{i,t} = \lambda_{i,t}^{½}\,\varepsilon_{i,t}$, where $\varepsilon_{i,t} \sim WN(0,1)$ and $i = 1,\ldots, p$. If $\lambda_{i,t} = \lambda_i\ \forall\ t$, this setup would not reproduce variance clusters since the conditional variance matrix will be time invariant. Since this feature is clearly detectable in our data, we will adopt the popular GARCH modeling in three alternative specifications. In mid eighties the Generalized ARCH or GARCH model was introduced (Bollerslev, 1986). Accordingly, the conditional volatility was modeled as

$$\lambda_{i,t} = \alpha_0 + \sum_{k=1}^{q}\alpha_k\varepsilon_{t-k}^2 + \sum_{j=1}^{p}\beta_j\lambda_{i,t-j}\,. \tag{3.6}$$

For a review of ARCH and GARCH models and their financial applications we refer the reader to the available monographs (Gourieroux, 1997; Franses and Van Dijk, 2000). A GARCH (1,1) model is equivalent to an infinite ARCH representation with exponentially declining weights on the lagged squared errors. In its original form, this process cannot reproduce important empirical evidence, such as persistence in variance changes, leverage and long memory, arising from the analysis of financial volatility. Consequently, we briefly review three evolutions of this model and their basic properties. For a more detailed review of these models and their Maximum Likelihood estimators, we refer the reader to Zivot and Wang (2003).

Non-stationarity. If the stationarity conditions do not hold, the process is said to be integrated GARCH or IGARCH. In the GARCH setting, non-stationarity has different meaning in comparison to the ARMA framework. For GARCH processes, when the sum of the coefficients is equal to one, the variance has unbounded support, whereas in ARMA models this implies unbounded support for the mean. Empirically, non-stationarity of GARCH processes causes high persistence of variance shocks. We can explore the GARCH non-stationarity substituting in equation (3.6) (when $p = q = 1$) $\alpha_1 = 1 - \beta_1$, so that we have

$$\lambda_{i,t} = \alpha_0 + (1-\beta_1)\varepsilon_{t-k}^2 + \beta_1\lambda_{i,t-1}\,, \tag{3.7}$$

for a smoothing coefficient $\beta_1 \in (0,1)$. By iterating the substitution, we have the interesting result

$$\lambda_{i,t} = \frac{\alpha_0}{(1-\beta_1)} + (1-\beta_1)(\varepsilon_{t-1}^2 + \beta_1 \varepsilon_{t-2}^2 + \beta_1^2 \varepsilon_{t-3}^2 + \beta_1^3 \varepsilon_{t-4}^2 + \ldots). \tag{3.8}$$

When $\alpha_0 = 0$ this is equivalent to an infinite Exponentially Weighted Moving Average (EWMA), such as the method used by J.P. Morgan in the RiskMetrics procedure, (see Longerstaey and Spencer, 1996). This parameterization reduces to a single, exponentially decaying coefficient that can be easily estimated even in a multivariate setting. This procedure has represented the benchmark approach for daily volatility forecasting.

Leverage and Heavy Tails. In classic GARCH models positive and negative shocks have the same effect on determining conditional variance, since this variable exclusively depends on squared residuals. However, a commonly observed fact in financial volatility is that bad news (negative shocks) tends to have a larger impact on volatility than good news (positive shocks). The Exponential GARCH (EGARCH) model has been proposed (Nelson, 1991) to allow for leverage effect. In this model the conditional variance evolves accordingly to

$$\tilde{\lambda}_{i,t} = \alpha_0 + \sum_{k=1}^{q} \alpha_k \frac{|\varepsilon_{t-k}| + \gamma_i \varepsilon_{t-1}}{\lambda_{i,t-1}} + \sum_{j=1}^{p} \beta_j \tilde{\lambda}_{i,t-j}, \tag{3.9}$$

where $\tilde{\lambda}_{i,t} = \log(\lambda_{i,t})$. The response variable is now the log-variance instead of the variance. This will ensure the positivity of the conditional variance without needing coefficients restrictions typically required by classic GARCH model. When ε_{t-k} is positive (or there is good news), the innovation effect is $(1 + \gamma_i)|\varepsilon_{t-1}|$; in contrast, when ε_{t-k} is negative (or there are bad news), the total effect is $(1 - \gamma_i)|\varepsilon_{t-1}|$. Bad news can now have a larger impact on the variance, and the value of γ_i would be expected to be negative. In the following analyses we will exclusively refer to EGARCH(1,1) models. In order to take into account the occurrence of large variance realizations, we will fit EGARCH models with t−distributed innovations.

Leverage and Long Memory. The EGARCH approach is very useful in order to provide a robust approach for modeling time series that are often encountered in finance. In fact, it can reproduce the so called *leverage effect* and a positive definite conditional variance, without needing to constrain model coefficients. Another important empirical feature characterizing financial volatility is the slow decay of its autocorrelation function. This phenomenon is the so called *long range dependence* also known as *long memory*, (see Mikosch and Stărică, 2003). Incorporating long memory and leverage into conditional volatility modeling lead to the direct extension of the Nelson's EGARCH (Bollerslev and Mikkelsen, 1996). In order to reproduce these features, a model has been proposed, where the conditional variance evolves accordingly to:

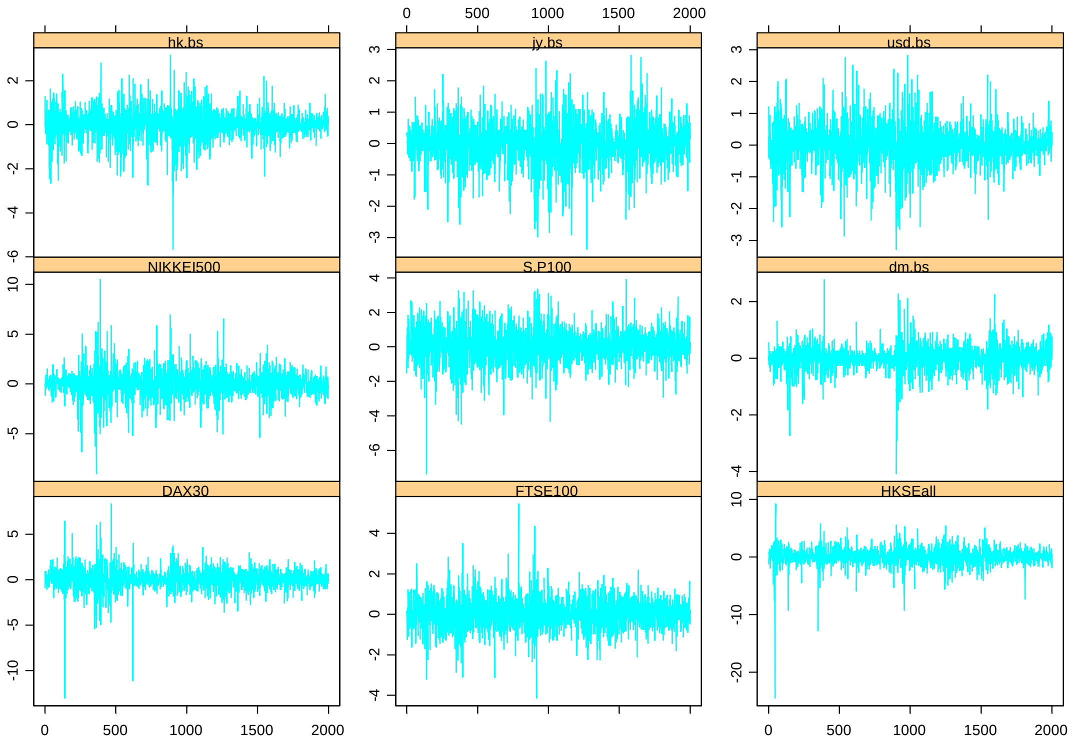
$$\phi(L)(1-L)^d \tilde{\lambda}_{i,t} = c + \sum_{j=1}^{q} \left(\tau_j |h_{i,t-j}| + \gamma_j h_{i,t-j}\right), \tag{3.10}$$

where L is the lag operator, $(1 - L)^d$ is the fractional difference operator (for $d < 1$), $\phi(L)$ is the stationary autoregressive polynomial and $h_{i,t} = \varepsilon_t/\lambda_{i,t}$ are the standardized residuals. Bollerslev and Mikkelsen (1996) proved that this model (that is commonly called Fractionally Integrated Exponential GARCH, or FIEGARCH, model) is stationary if $0 < d < 1$. In the following analyses we will exclusively refer to FIEGARCH (1,1) models.

4. ESTIMATING PORTFOLIOS

This section contains some statistical analyses which are intended to be preliminary to our main comparison based on economic criterions. Equation (2.1) shows that the optimal portfolio weights depend upon the VC matrix through its inverse. In practice the VC matrix is unknown and needs to be estimated from data. Since portfolio selection is mainly relevant as optimal investment prediction, it turns out that financial operators are interested in calculating the optimal weights relative to a future period. The length of this period typically depends upon the investment horizon. Investment Banks choose quite short horizons to recalibrate their portfolios, while Pension Funds are interested in longer horizons, since the purpose of their investment is typically be less speculative. The predictions are also determined by the methods and models used to estimate the VC matrix and the vector of expected returns. In the following sections we will assess and compare the performances of three models, evaluating portfolio sensitivity under both statistical and economic perspectives. We will adopt the common framework of OFV models, where the factors are assumed to evolve according to three alternative GARCH (1,1) specifications. We will first consider the EWMA model, deriving from the Integrated GARCH (IGARCH) model as illustrated in Section 3.1.

Figure 4.1 Log return series for five stock indices and four exchange rates observed in the period 4/4/1989 - 2/12/1996.



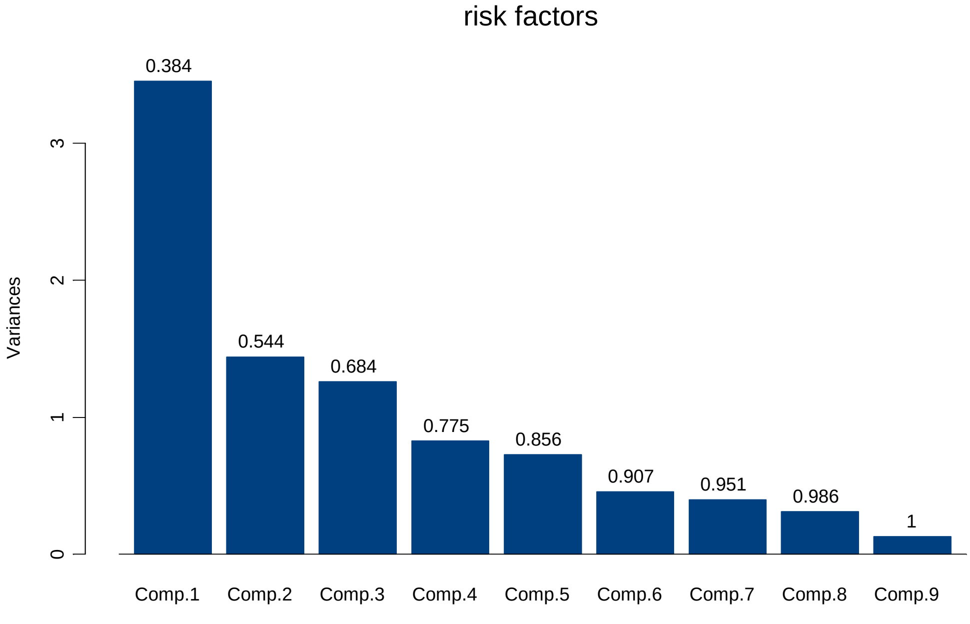
Notes: From the top row, clockwise: HKD/BP, JY/BP, USD/BP, NIKKEY 250, S&P 100, GM/BP, DAX 50, FTSE 100, HKSE all.

We will then consider the so called Exponential GARCH (EGARCH) model that is capable of reproducing the negative skewness in returns and factor’s distribution. Finally, we will consider the so called Fractionally Integrated Exponential GARCH (FIEGARCH) model that can also reproduce the long range dependence often observed in financial volatility. Orthogonal GARCH (OGARCH) models have been largely used by the financial industry for long time, and still provide a reliable instrument of wide practical application. Despite the availability of multivariate models (however, with somehow restricted parameterizations), OGARCH are still considered a robust framework that deserves to be further analyzed for its predictive ability. We have deliberately chosen a dataset of strongly correlated asset returns in order to mimic the way this modeling strategy has been adopted in the financial industry. The following analysis, therefore, should be of major interest to financial analysts, usually very interested in computationally robust methodologies.

Several authors note that by considering strongly correlated assets, the OFV models can provide a good approximation of the optimal factor determination. The choice of OFV models will allow us to compare relative merits of the aforementioned GARCH specifications, something that is not always possible to attain by considering multivariate GARCH, for which such extensions are often not available. In our study, log-returns of five stock indices (S&P 100, FTSE 100, NIKKEY 250, HKSE-all-shares and DAX 50) and four exchange rates (the exchange rate of British Pound (BP) vs. four currencies: US Dollar (USD), German Mark (GM), Japanese Yen (JY) and Hong Kong Dollar (HKD) have been considered in order to take into account the correlation between markets and related currencies.

In addition to statistical diagnostics, performances of different methods can be reasonably assessed looking at the consequences on management strategies. This comparative analysis will be conducted in Section 5, by fitting our models either on the fixed sample displayed in Figure 4.1, and on rolling windows starting with that sample. In either case, the sample and the windows will be of size 2000.

Figure 4.2 Scree plot for Orthogonal Variance Factors $y_1,\ldots, y_9$

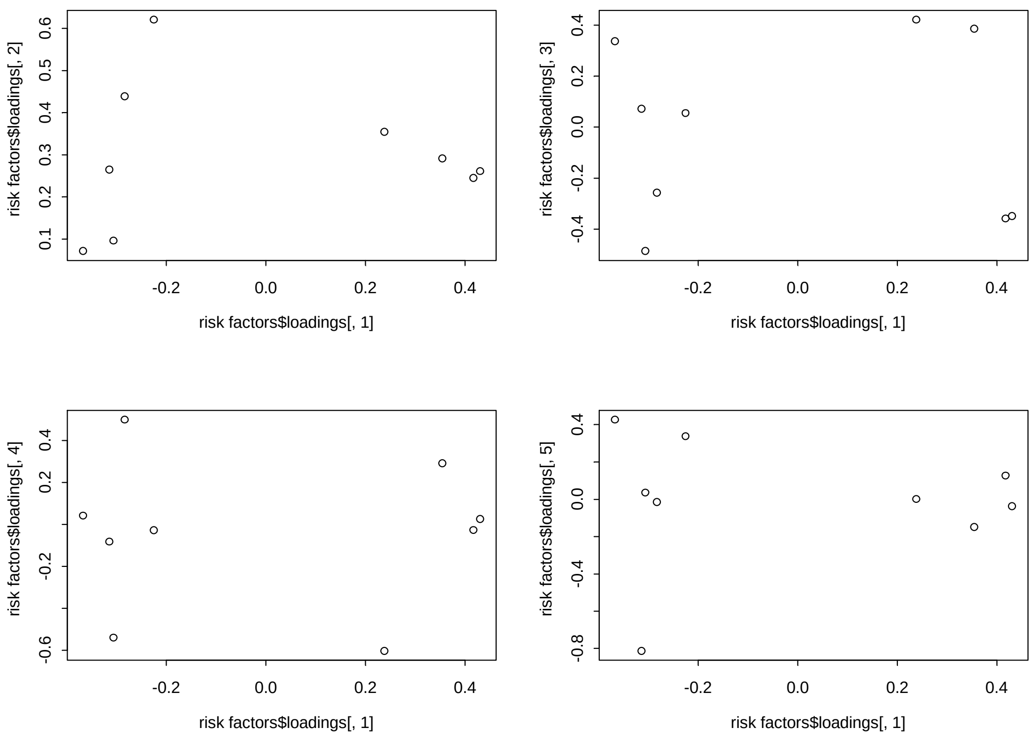


4.1. Estimating PCA

In this section we will estimate orthogonal factors for the dataset illustrated in Figure 4.1. In particular, we analyze the covariance matrix using PCA and represent graphically the most important loadings. We also look at the eigenanalysis of the sample covariance matrix, so that the cumulative eigenvalues can be represented using a simple scree-plot.

From the scree-plot (Figure 4.2) there is evidence that 5 principal components explain a large part of the overall variance. Figure 4.3 displays the corresponding principal components loadings. The first component explains a large part of the variance (about 38 %) and has been kept fixed. The other factors are plotted on the vertical axes. Figure 4.3 contains clear evidence that the first principal component can be interpreted as representing the negative correlation between market indices and currencies. The remaining principal components loadings are difficult to interpret.

Figure 4.3 Orthogonal Factors Loadings



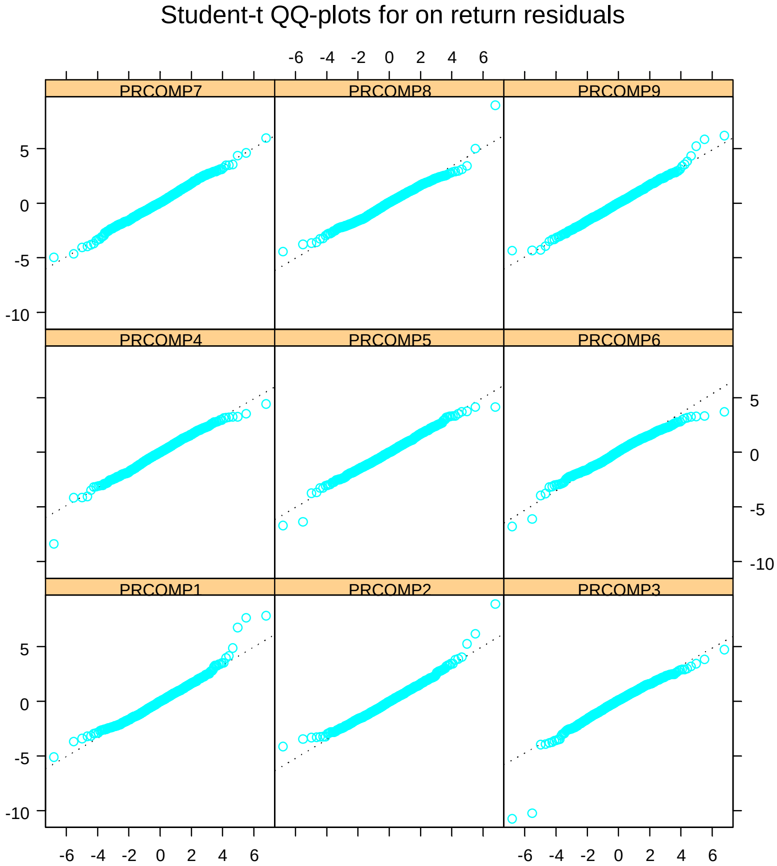
Notes: In the horizontal axis y_1 is kept fixed for all plots. In the vertical axis, from the top row, clockwise: y_2, y_3, y_4, y_5. We have evidence that y_1 represents the negative correlation between stock indices and currencies.

4.2. Comparing Fitted Models

The popular EWMA model, which has been widely used by the investment industry, in our comparative analysis serves as benchmark against which we assess, through the EGARCH platform, the relative merits of modeling heavy tails and long memory. In both latter cases, the EGARCH model has been chosen since the in-sample analysis has shown significant leverage effect. We recall that our main aim is to assess the various models through their portfolio performances. However, in this section we conduct a statistical in-sample analysis which is preliminary to the out-of sample comparison. Therefore, in this section we discuss several evidences arising from fitting the aforementioned models to orthogonal factors as estimated from our dataset. The first analysis aims at finding a convenient heavy tailed distribution for the EGARCH model to be used in the subsequent comparative study. We considered a Student-*t* distribution, with increasing degrees of freedom, to find the best fitting of EGARCH residuals. Figure 4.4 shows quantile plots for the best fitting distribution, which has six degrees of freedom. We then perform a further preliminary analysis by comparing the outcome of statistical diagnostics for the three factor models, where the fitted EGARCH assumes student-t distributed residuals.

We henceforth call this model the *t*-EGARCH model. We analyze residuals, obtained from fitting the three models, by testing the null hypotheses about the absence of serial correlation and ARCH effects. We use the Ljung-Box *portmanteau* test for the first purpose and the Breusch-Pagan test to conduct the second analysis. The results of the two tests are summarized in Table 4.1 and Table 4.2, respectively.

Figure 4.4 QQ plots between Gaussian EGARCH factor model residuals and Student-*t* random draws with 6 d.o.f.



Notes: From the bottom row, clockwise: $y_1,\ldots, y_9$:

MODEL	I		E		F	
FACTORS	LJ test	*p*.val	LJ test	*p*.val	LJ test	*p*.val
$y_{1,t}$	115.76	0.13	62.60	1	81.0	0.9
$y_{2,t}$	68.81	0.99	149.27	0	131.1	0.02
$y_{3,t}$	188.88	0	28.25	1	36.5	1
$y_{4,t}$	19.80	1	72.89	0.98	80.9	0.92
$y_{5,t}$	46.43	1	64.32	1	64.1	1
$y_{6,t}$	37.46	1	89.75	0.76	81.1	0.9
$y_{7,t}$	76.88	0.96	117.23	0.11	110.2	0.23
$y_{8,t}$	248.34	0	55.69	1	69.7	0.99
$y_{9,t}$	169.99	0	127.41	0.03	108.3	0.27

Table 4.1 Ljung-Box (LJ) test for the three orthogonal factor models: the EWMA-IGARCH model (I), the *t*–EGARCH (E) and FIEGARCH (F).

According to the tables above, none of the three factor models is remarkably superior to the other contenders. However, we have evidence that the FIEGARCH model is somehow superior to its competitors in terms of Ljung-Box test (absence of auto-correlation in residuals), whereas the EWMA-IGARCH model is somehow superior to its contenders in terms of Breusch-Pagan test (absence of ARCH effects in residuals).

MODEL	I		E		F	
FACTORS	BP test	*p*.val	BP test	*p*.val	BP test	*p*.val
$y_{1,t}$	7.93	1	16.18	1	68.30	0.99
$y_{2,t}$	9.65	1	1166.82	0	973.34	0
$y_{3,t}$	37.42	1	0.46	1	0.94	1
$y_{4,t}$	0.36	1	69.36	0.99	77.51	0.95
$y_{5,t}$	0.30	1	5.03	1	9.87	1
$y_{6,t}$	0.26	1	5.82	1	16.84	1
$y_{7,t}$	44.51	1	144.39	0	106.00	0.3
$y_{8,t}$	691.58	0	3.64	1	9.12	1
$y_{9,t}$	2.36	1	286.24	0	313.10	0

Table 4.2 Breusch-Pagan (BP) test for the two orthogonal factor models: *t*-EGARCH (E) and FIEGARCH (F).

5. COMPARATIVE OUT-OF-SAMPLE ANALYSIS

In this section we compare GARCH factor models in terms of out-of-sample portfolio performances. We will measure portfolio sensitivity to different approaches for predicting the conditional covariance matrix and the vector of expected returns. More precisely, the following criteria will be used:

- Efficient frontier analysis: this qualitative analysis will compare how different models produce out-of-sample efficient frontiers. Based on a fixed sample of 2000 observations, for each model four different forecasting horizons are considered (1 week, 2 weeks, 1 month, 2 months), in order to capture changing in performance due to different forecasting horizons. We will prefer models giving a stable estimation of portfolio mean and variance so that the frontier will not be too sensitive to the forecasting horizon.

- Turnover: this is an important criterion for asset management. In presence of transaction costs, asset sales or purchases are charged of a percentage that decreases net returns, so we will prefer a model that, given a certain one-step-ahead forecasting power, ensures more stable weights requiring a smaller amount of costs to update the efficient investment. Here we use a measure of absolute turnover given by:

$$\hat{\eta}_q = \frac{1}{500} \sum_{t=1}^{500} \mathbf{1}'(|\mathbf{w}_{t+1,t;q} - \mathbf{w}_{t,t-1;q}|), \tag{5.11}$$

  where $\mathbf{w}_{t+1,t;q}$ is a set of efficient weights relative to a risk appetite level q forecasted at time t for time $t+1$. As in the previous exercise, the backtest sample size is 500. A previous work where this assessment criterion was adopted is Gerhard and Hess (2003). The new aspect of our analysis is in that this measure is now calculated for different risk appetite levels, and a relation between costs (turnover) and risk can be evaluated for mean-variance portfolios.

- Forecasted residuals chart: We will use 500 time-points, ranging from (4/12/96) till (3/11/98), to build-up a backtest exercise on one-step-ahead standardized portfolio returns. The standardization is conducted by means of forecasted conditional portfolio mean and variances estimated from a rolling window of 2000 observations. For a model capable to capture the serial dependence of variance factors, we expect the standardized portfolio returns (using conditional mean and variance) to behave as independent and (Gaussian) identically distributed random errors. For each risk appetite level q, and for each forecasting horizon $\tau > 0$, we then measure the empirical size of Gaussian confidence regions (of 95% nominal level) built-up from these standardized residuals using the empirical measure:

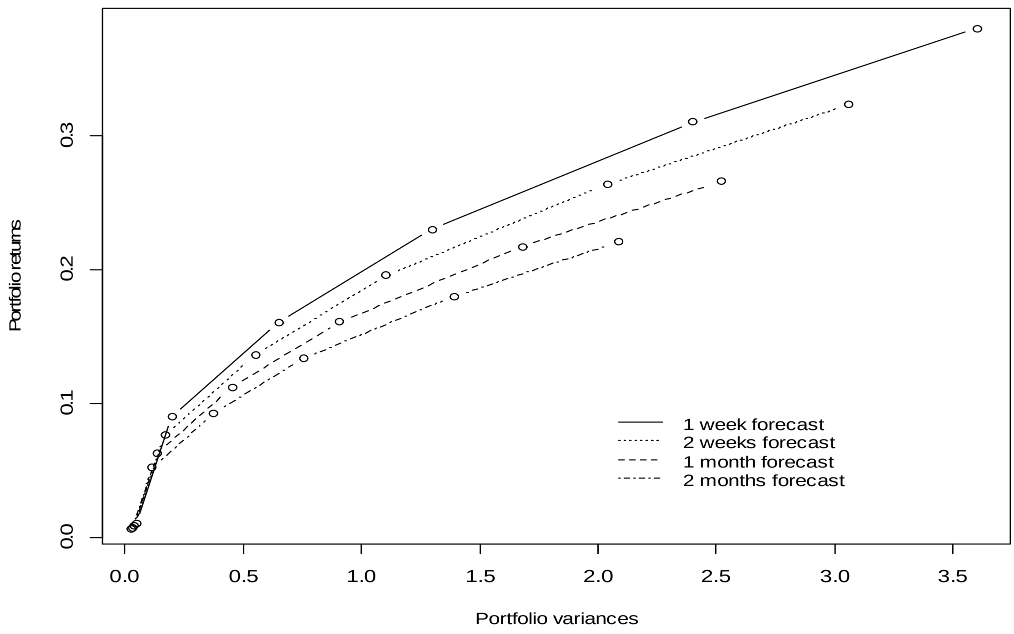
$$\hat{\theta}_{\tau;q} = \frac{1}{500} \sum_{t=1}^{500} \mathbb{I}\left(\left| \frac{\mu_{t+\tau;q} - m_{t,t+\tau;q}}{\sqrt{v_{t,t+\tau;q}}} \right| \leq 1.96 \right), \tag{5.12}$$

where $\mathbb{I}(\cdot)$ is the indicator function, $m_{t,t+\tau;q}$ is the portfolio (at time) t conditional mean forecasted for time $t+\tau$ and $v_{t,t+\tau;q}$ is the portfolio t conditional variance forecasted for time $t+\tau$. In this formula, each value of $t = 1,\ldots,500$ corresponds to a daily observation within the backtest period.

5.1. Efficient Frontier Portfolios

Looking at the entire frontier, we want to extend our comparative analysis in order to inspect how different forecasting models perform with respect to different risk levels. In this analysis we will consider forecasting horizons of 1 and 2 weeks and 1 and 2 months. Figure 5.5 shows the frontier estimated for the EWMA-IGARCH model.

Figure 5.5 Out-of-sample Efficient Frontier for EWMA Factor Model.



We have evidence that this model delivers portfolios with larger mean and variances, in comparison to the other models. The figure also shows a marked sensitivity of frontiers, with respect the forecasting horizon. The second frontier will be estimated by the Orthogonal EGARCH model, fitted using a Student-t distribution. Looking at Figure 5.6 we have evidence that, also in this case, both mean and variance estimations decrease for larger horizons, and their magnitude is directly related to the risk level. The portfolios appear to be more stable than in the previous case, and both means and variances lie in smaller ranges. Figure 5.7 shows the frontier estimated by using the FIEGARCH factor model. There is now a similar behavior to the previous model, although this strategy seems to be characterized by higher sensitivity to risk and forecasting horizon. Summarizing, all models showed decreasing estimates for both mean and variance in relation to the forecasting horizons. This decrease is more marked for the EWMA model, but it's also consistent for the FIEGARCH model. The t−EGARCH model seems to deliver the more stable forecasting, in relation to the forecasting horizons that we have considered.

Figure 5.6 Out-of-sample Efficient Frontier for EGARCH Factor Model**.**

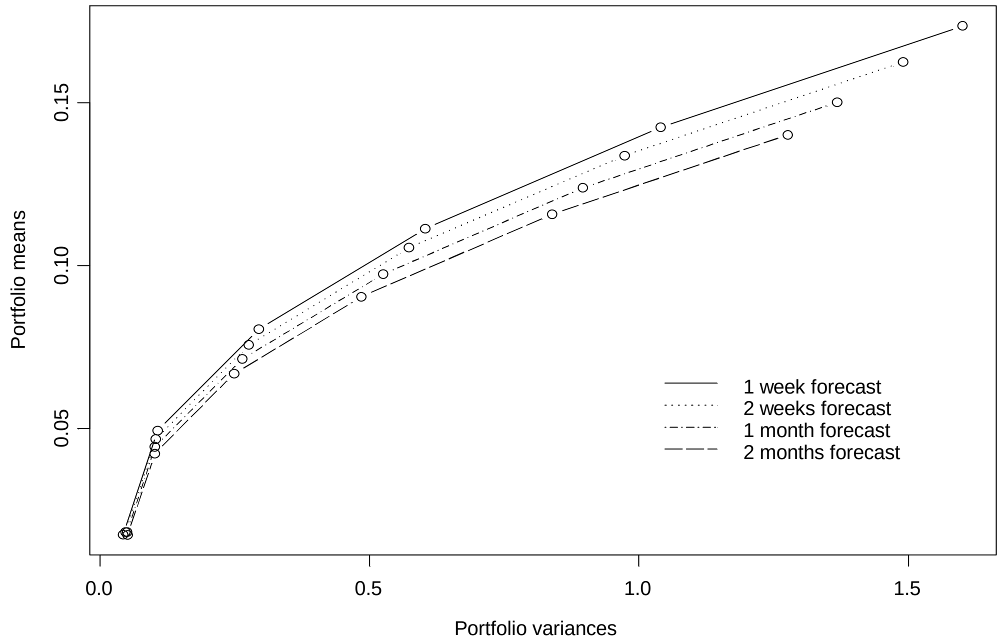
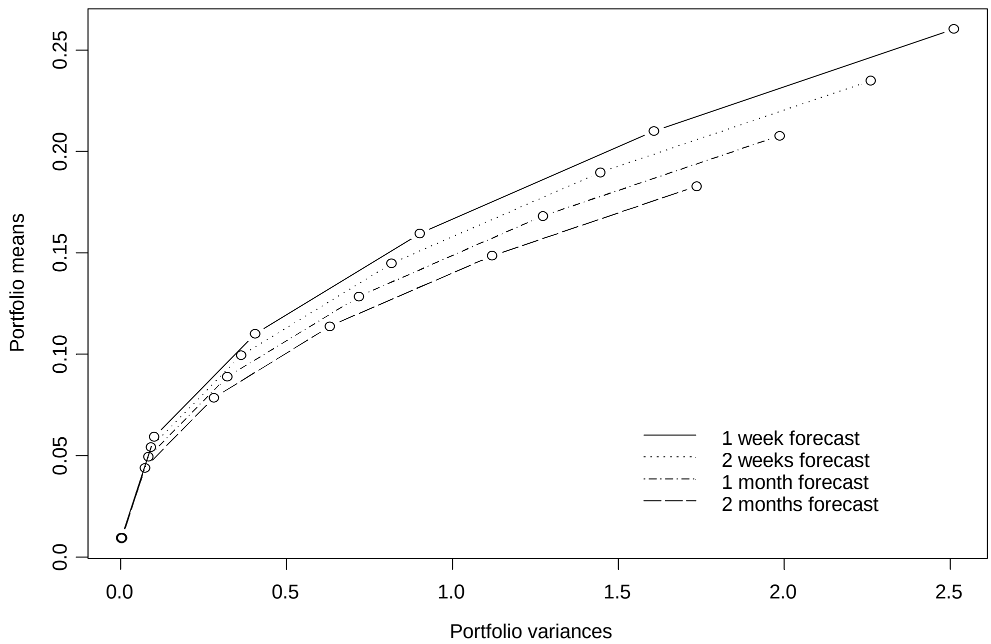


Figure 5.7 Out-of-sample Efficient Frontier for FIEGARCH Factor Model**.**



5.2. Portfolios Turnover

This criterion compares the stability of weights in a sequence of forecasted efficient portfolios. This measure is relevant in presence of transaction costs, when it represents the expected cost for maintaining an efficient investment. In the previous section we have formalized the measure used in this analysis. Table 5.3 displays the results of estimated turnovers using the three factor models.

The FIEGARCH model shows the lowest turnover when compared with the competitive models. The *t*-EGARCH model show lower turnover than the FIEGARCH model for minimum variance portfolios, but then increasingly underperform the FIEGARCH for increasing values of the risk appetite parameter. The EWMA-IGARCH uniformly underperforms both competitive models, except for minimum variance portfolios, for which it shows superior performance in comparison to both contenders. In particular, since minimum variance portfolio weights do not depend upon the expected portfolio returns, the EWMA model seems penalized by its inferior forecasting ability for this latter quantity.

q	EWMA	*t*-EGARCH	FIEGARCH
0	0.72	1.05	2.39
2	22.20	13.60	11.32
4	34.76	26.37	20.59
6	46.15	39.32	30.00
8	58.25	52.50	39.55
10	70.55	65.82	49.32

Table 5.3 Portfolio absolute turnover $\hat{\eta}_q$ for the three orthogonal factor models: EWMA (IGARCH), *t*-EGARCH and FIEGARCH.

5.3 Forecasting Standardized Portfolio Returns

In this analysis we have used up to 500 observations to construct out-of-sample charts of conditional standardized portfolio returns. Table 5.4 summarizes the results of this experiment, carried out in relation to different forecasting models and horizons.

q	E				E				F			
	1w.	2w.	1m.	2m.	1w.	2w.	1m.	2m.	1w.	2w.	1m.	2m.
0	.53	.53	.53	.53	.84	.83	.83	.81	.65	.64	.64	.63
2	.53	.53	.52	.52	.82	.81	.81	.79	.68	.68	.68	.67
4	.52	.52	.52	.52	.81	.80	.80	.78	.69	.69	.68	.68
6	.51	.52	.52	.52	.83	.81	.79	.77	.69	.68	.67	.67
8	.52	.52	.52	.52	.83	.81	.79	.78	.70	.70	.69	.67
10	.52	.52	.52	.52	.83	.81	.80	.78	.70	.69	.67	.65

Table 5.4 Empirical significance level of prediction intervals $\hat{\theta}_{\tau;q}$ for the EWMA-IGARCH (I) and the two orthogonal factor models: *t*-EGARCH (E) and FIEGARCH (F).

Looking at Table 5.4, the FIEGARCH model appears very robust with respect to risk appetite and to different forecasting horizons. However, the use of a Gaussian distribution causes the effective interval sizes to be sensibly lower than the theoretical level. The Orthogonal EGARCH model shows the best forecasting performance. The effective significance level is not much lower than the nominal level, and is very superior to the results obtained from the FIEGARCH model. Instead, the EWMA-IGARCH has the worst performance. Summarizing, we have evidence that both FIEGARCH and *t*-EGARCH provide reasonably good prediction intervals. The EWMA model provides less reliable forecasting, in relation to the considered forecasting horizons. This behavior seems to be due to its heavily constrained parameterization. A word is in order, about the sensitivity of these results upon the window length (2000 time points) that we have used. The stability of the empirical confidence levels across the forecasting horizons, and the simulation studies conducted in various studies (see Demos and Kyriakopoulou, 2010) suggest that this sample size is adequate to obtain reliable estimates of parameter values. In other words, the bias of estimates does not decrease significantly for increasing sample size. Bias corrections can always be applied, but in most cases, these provide some improvements for realistic sample sizes, of at most 3000 observations. In any cases, given the evidence of aforementioned simulation studies, it is prudent to use samples of at least 500 observations.

6. CONCLUSIONS

We conducted a comparative study on the mean-variance portfolio performances for three different orthogonal factor variance (OFV) models. The orthogonal factors were estimated by means of principal component analysis. We therefore considered univariate GARCH-type models for predicting variance factors. We initially considered the Orthogonal EWMA-IGARCH, as benchmark model. This is the generalization of the popular RiskMetrics procedure, since it allows the single orthogonal factors to be smoothed individually. We then compared this popular model with EGARCH factors, where the relative merits of heavy tails and long memory have been taken into account. We preliminarily evaluated these models in terms of statistical diagnostics and then compared by means of portfolio performances. These latter analyses included both qualitative and quantitative out-of-sample comparisons, conducted in order to assess portfolio performances for different risk levels. From a qualitative standpoint, we evaluated out-of-sample efficient mean variance frontiers, in order to assess the sensitivity of portfolios to risk appetite. For a quantitative analysis we conducted a backtest exercise with standardized portfolio returns chart, along with a comparison of portfolio absolute turnovers. It turned out that the overall best performance was provided by the EGARCH model, when fitted with a Student-*t* distribution with 6 degrees of freedom. Overall, the EGARCH model outperformed the FIEGARCH model, which also showed reasonable performances along with the EWMA-IGARCH model. Despite its wide use in the financial industry, this latter model produced poorer portfolio performances in comparison to the EGARCH-type factors. This shows that, along with heavy-tails, the leverage effect also plays a significant role in the prediction of efficient portfolios. We conclude that the out-of-sample performances of mean-variance portfolios are more positively affected by the use of heavy- tailed factors rather than the inclusion of long memory effects, an important role being played by the leverage effect. An **S+** library **portfoliOGARCH** containing the software used in our analyses is in preparation and will be available from the author upon request.

REFERENCES

Adcock, C.J. (2004). Estimating UK Factor Models Using The Multivariate Skew Normal Distribution. In *Linear Factor Models in Finance*, ed. S.E. Satchell and J. Knight. Butterworth Heinemann, Oxford and Burlington: 12-29.

Alexander, C. (2001). A Primer on the Orthogonal GARCH Model. *ICMA Centre Discussion Paper*. Available at *http://carolalexander.org/publish/download/DiscussionPapers.*

Best, M. and R. Grauer (1991). On the sensitivity of Mean-Variance Portfolios to changes in Asset Means: Some Analytical and Computational Results. *Review of Financial Studies*, 4: 315-342.

Bystrom, H.N.E. (2004). Orthogonal GARCH and covariance matrix forecasting: The Nordic stock markets during the Asian financial crisis 1997-1998. *European Journal of Finance*, 10(1): 44-67.

Bollerslev, T. (1986). Generalized Autoregressive Heteroskedasticity. *Journal of Econometrics*, 31: 307-327.

Bollerslev, T. and H.O. Mikkelsen (1996). Modeling and Pricing Long Memory in Stock Market Volatility. *Journal of Econometrics*, 73: 151-184.

Chopra, V. and W. Ziemba (1993). The Effect of Errors in Means, Variances and Covariances on Optimal Portfolio Choice. *Journal of Portfolio Management*, 19(2): 6-11.

Demos, A and D. Kyriakopoulou (2010). Bias Correction of ML and QML estimators in the EGARCH (1,1) Model. *DEOS Working Papers, Athens University of Economics and Business.*

Franses, H. and D. Van Dijk (2000). Non linear time series models in empirical finance. Cambridge: Cambridge Univ. Press.

Gerhard, F. and D. Hess (2003). Multivariate Market Risk Estimators: Reliability and Transaction Costs in the Context of Portfolio Selection. *European Journal of Finance*, 9(1): 1-18.

Gourieroux, C. (1997). *ARCH Models and Financial Applications*. Berlin: Springer Verlag.

Hawkes, R. and P. Date (2007). Medium-term horizon volatility forecasting: A comparative study. *Applied Stochastic Models in Business and Industry*, 23(6), 465-481.

Longerstaey, J., and M. Spencer (1996). RiskMetrics-Technical document. New York: *RiskMetrics Group, J.P.Morgan.*

Markowitz, H.M. (1952). Portfolio Selection. *The Journal of Finance* **7**(1): 77–91.

Mikosch, T. and C. Stărică (2003). Long range dependence effects and ARCH modeling. In *Theory and applications of long-range dependence*, ed. P. Doukhan, G. Oppenheim and M.S. Taqqu. Boston: Birkhäuser: 439-459.

Nelson, D.B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, 59(2): 347-370.

Zivot, E. and J. Wang (2006). Modeling Financial Time Series with S-PLUS, New York: Springer Verlag.