Minimum Variance Portfolio Selection for Large Number of Stocks – Application of Time-Varying Covariance Matrices

Abstract. An evaluation of the efficiency of different methods of the minimum variance portfolio selection was performed for seventy stocks from the Warsaw Stock Exchange. Eight specifications of multivariate GARCH models and six other methods were used. The application of all considered GARCH-class models was more efficient in stocks allocation than the implementation of the other analyzed methods. The simple specifications of multivariate GARCH models, whose parameters were estimated in two stages, like the DCC and CCC models were the best performing models.

Keywords: multivariate GARCH models, time-varying covariance matrix, portfolio selection.

Introduction

The selection of estimators of the population mean, variance and covariance of financial returns and closely connected with it the selection of forecasting methods of the population mean, variance and covariance of returns plays a vital role in the construction of efficient portfolios. There is a consensus among both financial market practitioners and scientists, that financial returns are difficult to forecast and the portfolio construction process according to the criteria proposed by Markowitz (1952, 1959) is very sensitive to the selection of an estimator of the expected value of returns (see Michaud, 1989; Best, Grauer, 1991; Chopra, Ziemba, 1993). Small differences in estimates of the expected returns often lead to a meaningful portfolio reconstruction (see e.g. Jobson, Korkie, 1980). It is believed that the estimation of population variances and covariances of returns is easier than the estimation of population means (Merton 1980; Nelson, 1992), however the selection of estimators of population variances and covariances of returns has also a significant impact on an asset allocation (see e.g. Litterman, Winkelmann, 1998; Chan, Karceski, Lakonishok, 1999). Traditionally used estimators of the population mean, variance and co-
variance, namely the arithmetic mean and empirical variance and covariance calculated on the basis of available historical data do not give the best results in efficient portfolios construction (see e.g. Jorion, 1991; Sheedy, Trevor, Wood, 1999; Johannes, Polson, Stroud, 2002; Flavin, Wickens, 2006).

The classical approach to the selection of efficient portfolios is static, i.e. a chosen period or moment is considered, but not their sequence. Such an approach ignores therefore the variability of conditional variances and covariances of returns. A dynamic approach to the selection of efficient portfolios for homogeneous assets, i.e. stocks, based on the forecasts of variances and covariances of returns constructed from multivariate GARCH models is presented in this paper. There is an extensive literature on time-varying conditional variances of returns but also on time-varying conditional covariances and correlation coefficients between financial series. The multivariate GARCH process allows to describe both time-varying conditional variances and covariances of returns. If variances and covariances of returns are not constant, then the forecasts based on multivariate GARCH models should give additional benefits in the selection of efficient portfolios. The literature on the application of GARCH models in construction of portfolios is poor, and the results of such analyses, due to the complexity of the problem, are still fragmentary. In most of the studies the univariate GARCH model is applied or a very limited number of assets are used. Among the few investigations in which multivariate GARCH models were applied for large portfolios the following papers can be mentioned: Engle, Sheppard (2001), Engle, Colacito (2006), Osiewalski, Pajor (2010).

The primary purpose of this study is to evaluate the effectiveness of different methods of the minimum variance portfolio selection, mainly with the application of various multivariate specifications of GARCH models. The adopted approaches to the portfolio construction differ only by forecasting methods of the covariance matrix of returns. The paper is an extension of the author's earlier work presented in Fiszeder (2004, 2007). The plan for the rest of the paper is as follows. Section 1 outlines the way in which efficient portfolios are selected. In section 2 competing methods of the covariance matrix estimation are presented. Section 3 contains the effectiveness analysis of the minimum variance portfolio selection for seventy stocks quoted on the Warsaw Stock Exchange (WSE) and section 4 presents the conclusions.

1. Dynamic Process of Portfolio Selection with Application of GARCH Models

It is difficult to evaluate the influence of the choice of the covariance matrix estimator against the selection of the population mean estimator in construction of efficient portfolios. One of the ways to eliminate the influence of the choice of the mean estimator is the minimum variance portfolio selection. Shares of
individual assets in the minimum variance portfolio depend solely on the covariance matrix, that is why such procedure is used in the paper.

For a given $t$, based on all the available data from the period $[1,t]$ parameters of a GARCH model are estimated\(^1\). The forecast of the covariance matrix based on the estimated model is formulated at time $t + \tau$, where the $\tau$ is the forecast horizon. The constructed forecast is used for the selection of efficient portfolio. Let $W_{t+\tau|\mathcal{F}_t} = (w_{t+\tau|\mathcal{F}_t 1}, w_{t+\tau|\mathcal{F}_t 2}, \ldots, w_{t+\tau|\mathcal{F}_t N})$, where $w_{t+\tau|\mathcal{F}_t i}$ is the share of asset $i$ in the portfolio at time $t + \tau$, $H_{t+\tau|\mathcal{F}_t}$ is the forecast of the conditional covariance matrix of returns at time $t + \tau$. The variance of portfolio returns is then equal to $W_{t+\tau|\mathcal{F}_t} H_{t+\tau|\mathcal{F}_t} W_{t+\tau|\mathcal{F}_t}$. In order to select the minimum variance portfolio (the global minimum) the following quadratic programming problem has to be solved:

$$W_{t+\tau|\mathcal{F}_t} H_{t+\tau|\mathcal{F}_t} W_{t+\tau|\mathcal{F}_t} \rightarrow \text{min},$$  

subject to the constraint:

$$W_{t+\tau|\mathcal{F}_t} 1 = 1,$$

where $1$ is the $N \times 1$ vector of ones.

When short selling is not allowed the boundary conditions have to be additionally imposed:

$$w_{t+\tau|\mathcal{F}_t i} \geq 0 \quad \text{for } i = 1, 2, \ldots, N.$$  

The whole procedure is repeated for successive periods (with the arrival of subsequent data).

If short selling is allowed, then shares of assets in the minimum variance portfolio are defined by the following formula:

$$W_{t+\tau|\mathcal{F}_t} = \frac{1}{C_{t+\tau|\mathcal{F}_t}} H_{t+\tau|\mathcal{F}_t}^{-1} 1,$$

where $C_{t+\tau|\mathcal{F}_t} = 1'H_{t+\tau|\mathcal{F}_t}^{-1} 1$.

The variance of the minimum variance portfolio is then equal to $V_{t+\tau|\mathcal{F}_t} = 1/C_{t+\tau|\mathcal{F}_t}$.

Other efficient portfolios can be selected by finding the minimum variance portfolio subject to a minimum expected return.

\(^1\) If conditional expected values of returns are different from zero, then parameters of conditional mean equations should also be estimated.
2. Specifications of Multivariate GARCH Models

Eight parameterizations of multivariate GARCH models were applied in the analysis: Scalar BEKK, Integrated, CCC, Orthogonal, DCC, Integrated DCC, DECO-DCC and additionally Scalar BEKK with Student-t innovations. The results obtained for the GARCH models were compared with the outcomes for six other methods: equal shares in all stocks, the unconditional covariance matrix of returns, the rolling covariance matrix\(^2\), the 25-day rolling covariance matrix, the exponentially weighted moving average estimate of the covariance matrix (hereafter the EWMA covariance matrix)\(^3\), the EWMA covariance matrix with a smoothing parameter set to 0.94 (hereafter the RiskMetrics\(^4\)).

Only some basic information about the considered multivariate models is given below. More details can be found for example in Bauwens, Laurent, Rombouts (2006) or Silvennoinen, Teräsvirta (2009).

Let assume that \(\varepsilon_i\) can be either returns with the mean zero or residuals from filtered time series:

\[
\varepsilon_i \mid \psi_{t-1} \sim N(0, H_t),
\]

where \(\psi_{t-1}\) is the set of all information available at time \(t-1\) and \(H_t\) is the \(N \times N\) symmetric conditional covariance matrix.

The estimation of parameters of the general form of a multivariate GARCH model, VECH model\(^5\) (Kraft, Engle, 1983), is very difficult even for a small number of assets. For this reason simpler parameterizations of multivariate GARCH models were used in the study. Considered were only those specifications which ensure the positive definiteness of the covariance matrix.

Baba, Engle, Kraft, Kroner (1990) introduced the following form the so-called BEKK(p,q) model\(^6\) (see Engle, Kroner, 1995):

\[
H_t = CC' + \sum_{i=1}^{q} D_i \varepsilon_{t-i} \varepsilon'_{t-i} D_i' + \sum_{j=1}^{p} E_j H_{t-j} E_j',
\]

where \(C, D_i, E_j\) are \(N \times N\) parameter matrices and \(C\) is an upper triangular matrix.

\(^2\) The value of a rolling window was chosen to minimize the variance of a portfolio in a pre-sample.
\(^3\) The value of a smoothing parameter was chosen to minimize the variance of a portfolio in a pre-sample.
\(^4\) The RiskMetrics methodology was developed in the investment bank J. P. Morgan for measuring market risk with VaR. The value of a smoothing parameter in the EWMA model was often set by financial market practitioners to 0.94 for daily data.
\(^5\) The name of the model comes from the application of the vech operator for the conditional covariance matrix.
\(^6\) The name of the model is formed by the first letters of the authors surnames.
The BEKK model is also too complex for large portfolios and the imposition of restrictions on parameters in estimation is necessary. The so-called Scalar BEKK model can be obtained by replacing the matrices $D_i$ and $E_j$ by the scalars $d_i^{1/2}$ and $e_j^{1/2}$. Additionally the variance targeting approach (see Engle, Mezrich, 1996) was used. For example for $q = p = 1$ instead of the product $CC'$ the following formula can be substituted:

$$CC' = (1 - d_i - e_j)S,$$

(7)

where $S$ is the sample covariance matrix given as $S = \frac{1}{T} \sum_{t=1}^{T} \epsilon_t \epsilon_t'$. Two specifications of the conditional distribution of $\epsilon_t$ in (5) were considered for the Scalar BEKK model, namely multivariate normal and Student-t.

A further simplification of the model can be obtained by the assumptions $CC' = 0$ and $\sum_{i=1}^{q} d_i + \sum_{j=1}^{p} e_j = 1$. This formulation is often called the Integrated multivariate model\textsuperscript{7}. For $q = p = 1$ the model has only one parameter.

In the Constant Conditional Correlations (CCC) model of Bollerslev (1990), which is outside the BEKK class, the time-varying conditional covariances are parameterized to be proportional to the product of the corresponding conditional standard deviations:

$$H_t = D_t \Gamma D_t',$$

(8)

where $D_t$ is a $N \times N$ diagonal matrix $D_t = \text{diag}(h_{t1}^{1/2}, h_{t2}^{1/2}, ..., h_{tN}^{1/2})$ with $h_{it}$ defined as any univariate GARCH model, and $\Gamma$ is a $N \times N$ matrix of the time–invariant conditional correlations.

Alexander and Chibumba (1996) introduced the Orthogonal GARCH model defined as:

$$V^{-1/2} \epsilon_t = u_t = \Lambda_m f_t,$$

(9)

$$H_t = V^{1/2} V^{1/2} f_t,$$

(10)

where $V = \text{diag}(v_1, v_2, ..., v_N)$, with $v_j$ the population variance of $\epsilon_{it}$, $\Lambda_m$ is a matrix of dimension $N \times m$ given by $\Lambda_m = P_m \text{diag}(l_1^{1/2}, l_2^{1/2}, ..., l_m^{1/2})$, $l_1 \geq l_2 \geq ... \geq l_m > 0$ being the $m$ largest eigenvalues of the population correlation matrix of $u_t$, and $P_m$ the $N \times m$ matrix of associated (mutually orthogonal) eigenvectors, $f_t = (f_{1t}, f_{2t}, ..., f_{mt})'$ is a random process such that $E_{t-1}(f_t) = 0$, the

\textsuperscript{7} The name of the model comes from an analogous parameterization of the univariate IGARCH model.
conditional covariance matrix of \( f_t \) is equal to \( Q_t = \text{diag}(\sigma_{f_t}^2, \sigma_{f_t}^2, \ldots, \sigma_{f_t}^2) \), \( \sigma_{f_t}^2 \) is defined as a univariate GARCH model (for \( i = 1,2,\ldots,m \)) and the conditional covariance matrix of \( u_t \) is equal to \( V_t = \Lambda_m Q_t \Lambda_m' \).

The Dynamic Conditional Correlation (DCC) model of Engle (2002) can be defined as:

\[
H_t = D_t R_t D_t', \quad (11)
\]
\[
R_t = Q_t^{-1} \Lambda_t Q_t'^{-1}, \quad (12)
\]
\[
Q_t = (1 - \sum_{i=1}^{O} \alpha_i - \sum_{j=1}^{P} \beta_j) S + \sum_{i=1}^{O} \alpha_i (\epsilon_{t-i} \epsilon_{t-i}') + \sum_{j=1}^{P} \beta_j Q_{t-j}, \quad (13)
\]

where \( D_t = \text{diag}(h_t^{1/2}, h_t^{1/2}, \ldots, h_t^{1/2}) \), \( h_t \) can be defined as a univariate GARCH model (for \( k = 1,2,\ldots,N \)), \( z_t \) is a vector of standardized values of \( \epsilon_{t-i} \), i.e. \( z_t = \epsilon_t / \sqrt{h_t} \), \( R_t \) is the time-varying conditional correlation matrix of \( z_t \), \( S \) is the sample covariance matrix of \( z_t \), \( Q_t' \) is a diagonal matrix composed of the square root of the diagonal elements of \( Q_t \) and the parameters have to satisfy the condition \( 0 < \sum_{i=1}^{O} \alpha_i + \sum_{j=1}^{P} \beta_j < 1 \).

Besides the standard DCC model two other modifications were also considered. When the sum of the parameters is equal to one \( \sum_{i=1}^{O} \alpha_i + \sum_{j=1}^{P} \beta_j = 1 \) the DCC model is called the Integrated DCC model. Engle and Kelly (2008) introduced the DECO-DCC (dynamic equicorrelation) model. In this parameterization of the DCC model the equality of all pairwise conditional correlations at each time is assumed.

3. Evaluation of Portfolio Performance for Seventy Stocks Quoted on the WSE

The presented approach for the minimum variance portfolio selection was evaluated for Polish stocks. The investigated period was November 17, 2000\(^8\) to June 30, 2009 (2158 daily returns). All stocks quoted in the specified period on the Warsaw Stock Exchange were considered in the analysis. The companies for which the percent of non-trading days was higher than 5% were omitted in order to avoid the problem of non-synchronous trading. In total seventy stocks were analyzed, however all the models used in the paper can be applied for

\(^8\) Since the introduction of the new trading system Warset.
a much larger number of companies. An evaluation of portfolio performance was based on data from January 2004 to June 2009 (1380 observations). The following five steps were performed: (1) the estimation of parameters of all considered models\(^9\) (at the beginning for data from November 17, 2000 to December 30, 2003)\(^10\), (2) the construction of one-day ahead forecasts of the conditional covariance matrix, (3) the minimum variance portfolio selection (4) the ex post calculation of the portfolio variance as a square of the realized portfolio return, (5) the extension of the sample with one observation. All the steps were repeated 1380 times. Every time the estimation of parameters was performed for increasing sample size. For each model and method the mean of the portfolio variances was calculated. The results are presented in Table 1 (a square root of the mean is given).

Table 1. Estimates of the standard deviations of returns for the minimum variance portfolios

<table>
<thead>
<tr>
<th>Portfolio designation</th>
<th>Standard deviation ((\times 10^{-2}))</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrated DCC</td>
<td>0.9257</td>
<td>1</td>
</tr>
<tr>
<td>DCC</td>
<td>0.9316</td>
<td>2</td>
</tr>
<tr>
<td>CCC</td>
<td>0.9326</td>
<td>3</td>
</tr>
<tr>
<td>DECO-DCC</td>
<td>0.9843</td>
<td>4</td>
</tr>
<tr>
<td>Scalar BEKK Student-t</td>
<td>1.0030</td>
<td>5</td>
</tr>
<tr>
<td>Scalar BEKK</td>
<td>1.0065</td>
<td>6</td>
</tr>
<tr>
<td>Orthogonal 70 factors</td>
<td>1.0854</td>
<td>7</td>
</tr>
<tr>
<td>Integrated</td>
<td>1.1155</td>
<td>8</td>
</tr>
<tr>
<td>Unconditional covariance matrix</td>
<td>1.1155</td>
<td>8</td>
</tr>
<tr>
<td>Rolling covariance matrix</td>
<td>1.2634</td>
<td>10</td>
</tr>
<tr>
<td>Equal shares in all stocks</td>
<td>1.3538</td>
<td>11</td>
</tr>
<tr>
<td>RiskMetrics</td>
<td>1.6549</td>
<td>12</td>
</tr>
<tr>
<td>EWMA covariance matrix</td>
<td>2.9846</td>
<td>13</td>
</tr>
<tr>
<td>25-day rolling covariance matrix</td>
<td>29.7629</td>
<td>14</td>
</tr>
</tbody>
</table>

The best performing model in the selection of the minimum variance portfolio was the Integrated DCC model. On the other hand the worst performing method was the 25-day rolling covariance matrix.

An application of all the GARCH-class models was more effective in allocation of stocks than the implementation of the other analyzed methods. The first four positions in the ranking (see Table 1) were occupied by the models, whose parameters were estimated in two steps and in the first stage parameters of a univariate GARCH model were estimated.

The Integrated model had the worst rank among all the applied GARCH models, however in this case the model was reduced to the unconditional covar-

---

\(^9\) Codes written by the author in the Gauss programming language were applied.

\(^10\) Logarithmic returns were used in the study. Because daily data were used, the logarithmic return on a portfolio is very close to the weighted average of the logarithmic returns on the individual assets.
The Orthogonal GARCH model, which assumes that returns depend on common independent factors extracted by principal component analysis ranked the one before last among the GARCH models. This result follows, first of all, from the loss of important information.

The methods used by financial market practitioners, i.e. the rolling covariance matrix and the exponentially weighted moving average estimate of the covariance matrix took a distant positions in the ranking, even behind the unconditional covariance matrix of returns. The very poor result of the rolling covariance matrix with a low value of the rolling window equal to 25 deserves to be highlighted. The value of the smoothing parameter, set often for daily data in the RiskMetrics methodology to 0.94 is not the optimal value for the Polish stock market (it is definitely too low).

For a large number of assets like in this study, probably none of the restrictions, assumed by the considered GARCH models, is met. However, one can attempt to indicate the restrictions, which have a stronger negative influence on the performance of portfolio, in the case when they are not met. For example, the assumption that models describing conditional variances and covariances have the same parameters for all series\(^{12}\) (like in the Scalar BEKK and Integrated models) is probably too strong. It also seems that the presupposition of the constancy of conditional correlation coefficients is less restrictive.

The statistical significance of the observed differences between the performance of the different models was not verified, however some of them are relevant from the economic point of view. For example, the difference between the first in the ranking, namely the Integrated DCC model and the unconditional covariance matrix of returns means a decrease of about 3 percentage points of the standard deviation per year.

Additionally, values of the Schwarz information criterion (SIC) were calculated, when it was possible to evaluate a joint likelihood function (see Table 2). The ranking of the models according to the SIC is similar to the one for the portfolio performance evaluation (Table 1). The exceptions are the Scalar BEKK model with Student-t innovations and the Integrated model which take better position in the SIC ranking. The better performance of GARCH models with Student-t innovations in the rankings based on the information criteria is common in other studies (see Fiszeder, 2009). The rankings of models constructed on the basis of the information criteria provide useful clues for the selection of models for the minimum variance portfolio construction. It has to be remembered, however, that some characteristics are crucial in the evaluation of the general fit of the model in a sample (like for instance a type of conditional density), but their influence on the performance of portfolio is not so important.

---

\(^{11}\) Estimates of the parameter on the lagged conditional covariance matrix were equal to one.

\(^{12}\) It means, that the volatility dynamics of all series is very similar.
Table 2. Ranking of the models based on the Schwarz information criterion

<table>
<thead>
<tr>
<th>Portfolio designation</th>
<th>SIC</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrated DCC</td>
<td>-679060.02</td>
<td>1</td>
</tr>
<tr>
<td>DCC</td>
<td>-666933.43</td>
<td>2</td>
</tr>
<tr>
<td>Scalar BEKK Student-t</td>
<td>-661928.17</td>
<td>3</td>
</tr>
<tr>
<td>DECO-DCC</td>
<td>-658854.85</td>
<td>4</td>
</tr>
<tr>
<td>Integrated</td>
<td>-656773.56</td>
<td>5</td>
</tr>
<tr>
<td>Scalar BEKK</td>
<td>-645459.24</td>
<td>6</td>
</tr>
<tr>
<td>RiskMetrics</td>
<td>2052174.1</td>
<td>7</td>
</tr>
</tbody>
</table>

It is interesting to compare the results of this analysis with the similar study for twenty stocks (Fiszeder, 2007). The GARCH models, whose parameters were estimated in one stage (like the Scalar BEKK and Integrated models) took further positions in the ranking. It seems that the assumption that models describing conditional variances have the same parameters for all series becomes more and more restrictive for a larger number of assets. Furthermore, the differences between evaluations of portfolios, obtained with the use of the applied methods, were greater for the seventy stocks. The reason for such results is probably the higher differentiation of the seventy stocks, among which are both huge but also very small companies. In the study for the twenty stocks only the biggest companies were considered. This comparison for different numbers of stocks clearly shows that the received results depend on the properties of financial time series.

Conclusions

The dynamic approach to the selection of efficient portfolios for a large number of homogeneous assets, i.e. stocks, based on the forecasts of variances and covariances of returns constructed from multivariate GARCH models has been presented in this paper. An evaluation of the efficiency for different methods of the minimum variance portfolio selection was performed for the seventy stocks from the Warsaw Stock Exchange. The eight specifications of multivariate GARCH models and the six other methods were used.

Capturing time-varying variances and covariances of stock returns does not always increase the efficiency of the asset allocation process. The application of all the considered GARCH-class models was more efficacious in the allocation of stocks than the implementation of the other analyzed methods, including the methods employed often by financial market practitioners. The simple specifications of multivariate GARCH models, whose parameters were estimated in two stages, like the DCC and CCC models, were the best performing models.
References


**Konstrukcja portfeli o minimalnej wariancji dla dużej liczby spółek – zastosowanie zmiennych w czasie macierzy kowariancji**

Zarys treści. W pracy dokonano oceny efektywności różnych metod tworzenia portfeli o minimalnej wariancji, w tym przede wszystkim z wykorzystaniem różnych specyfikacji wielorównaniowych modeli GARCH. Badanie zostało przeprowadzone dla 70 spółek notowanych na GPW w Warszawie. Zastosowano osiem parametryzacji modelu GARCH: skalarny BEKK, zintegrowany, CCC, ortogonalny dla 70 czynników, DCC; zintegrowany DCC, DECO-DCC, skalarny BEKK z warunkowym rozkładem t Studenta oraz sześć innych metod: równe udziały dla wszystkich aktywów, bezwarunkowa macierz kowariancji stóp zwrotu, ruchoma macierz kowariancji ze stałą wygładzania równej 25, metoda wyrównywania wykładniczego dla macierzy kowariancji oraz metoda wyrównywania wykładniczego dla macierzy kowariancji z parametrem wygasania równym 0,94.

Słowa kluczowe: wielorównaniowe modele GARCH, zmienająca się w czasie macierz kowariancji, konstrukcja portfela.