Volatility Forecasts and Value at Risk Evaluation for the MSCI North America Index

Momtchil Pojarliev[1] and Wolfgang Polasek[2]

[1] INVESCO Asset Management,
Bleichstrasse 60-62, D-60313 Frankfurt, Germany
[2] Institute of Advanced Studies, Vienna, Stumpergasse 56, 1060 Wien
Email: polasek@ihs.ac.at

Abstract. This paper compares different models for volatility forecasts with respect to the value at risk performance (VaR). The VaR measures the potential loss of a portfolio for the next period at a given significance level. We focus on the question if the choice of the appropriate volatility forecasting model is important for the VaR estimation. We compare the forecasting performance of several volatility models for the returns of the MSCI North America index. The resulting VaR estimators are evaluated by comparing the empirical failure rate with the forecasting performance.

1 Introduction

In April 2000 the NASDAQ 100 index dropped by about 20%. Similar developments in other stock markets have led to more research in risk assessments: Investors are interested in reliable risk measures of their portfolios, like the VaR measure, the lower quantiles of the predicted returns distribution. There are numerous approaches to calculate the VaR. Many analysts use the popular variance-covariance approach by J.P. Morgan, better known as RiskMetrics (1996). Recently, new methods have been developed on the basis on extreme value theory (Danielson and Vries (1997))and a comparative approach for Germany can be found in Zucchini and Neumann (2001). The aim of the paper is to investigate the following question: Can time series forecasts improve the VaR estimate and what evaluation criteria for volatility forecasts are good diagnostics for the VaR performance? Therefore we will compare the performance of the following models: The naive model, where the variance estimator is just the historical variance, the RiskMetrics model, the GARCH, an asymmetric GARCH (AGARCH) and a bivariate BEKK model.

The paper is organized as follows: In the next section we estimate the volatility of the MSCI North America index with the above models. The forecasting performance of the models is compared in section 3 using a rolling sample of 800 observation and an evaluation period of approximately two years. As forecasting evaluation criteria we use an auxiliary linear regression

model (Pagan and Schwert (1990)) and the Christoffersen tests (Christoffersen (1998)) for VaR. In section 4 we estimate the VaR of a hypothetical portfolio of 1 Mio \$ and compare the results. In the last section we conclude.

2 Volatility models

We will investigate the volatility of the daily returns of the MSCI North America index from May 1st 1995 until April 3rd 2000. The first 800 observations (from 1st May 1995 until 22 May 1998) are used as training sample and the remaining observation are used for out-of-sample comparison.

2.1 The naive model

The naive model uses the variance of a moving sample of 800 observation (approximately 3 years) as forecast for the next period (1000 observations were used in Dockner and Scheicher (1999)).

$$\hat{\sigma}^2_{t+1} = \frac{1}{799}\sum_{i=1}^{800}(r_{t+1-i} - \hat{\mu})^2 \tag{1}$$

where r_t are the returns at time t and $\hat{\mu}$ is the estimated average return of the sample.

2.2 The RiskMetrics model

The model proposed by J. P. Morgan (1995) is an exponentially weighted moving average model. The volatility of the next period can be calculated as a weighted average of the current volatility and the squared returns:

$$\sigma^2_{t+1} = \lambda\sigma^2_t + (1-\lambda)r^2_t \tag{2}$$

where λ is the weight factor. As proposed by RiskMetrics we set λ equal to 0.94. As an initial value for σ^2, we use the squared returns.

2.3 The GARCH model

The simplest form of observation model we can use for a returns vector is:

$$r_t = \mu + \epsilon_t, \tag{3}$$

where the conditional distribution of ϵ_t is assumed to be Gaussian with mean zero and variance σ^2_t. A GARCH(p,q) model uses the following parameterisation for the variance:

$$\sigma^2_t = \alpha_0 + \sum_{i=1}^{p}\alpha_i\epsilon^2_{t-i} + \sum_{i=1}^{q}\beta_i\sigma^2_{t-i}. \tag{4}$$

The information criteria AIC and BIC in Table 1 suggest to select a GARCH(1,1) model for the period from 1st May 1995 until 22 May 1998 for the daily returns of the MSCI North America index.

The estimated GARCH(1,1) model of the returns of the MSCI North America index is (t-values in parentheses):

$$\begin{array}{lllll} \hat{\sigma_t^2} & = 10^{-6}1.23 & + 0.083\epsilon_{t-1}^2 & + 0.904\hat{\sigma}_{t-1}^2. \\ (t-val.) & (2.48) & (7.80) & (57.47) \end{array} \quad (5)$$

2.4 The asymmetric GARCH model

An asymmetric GARCH(p,q) model (AGARCH) has the form

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p}(\alpha_i + \gamma_i S_{t-i})\epsilon_{t-i}^2 + \sum_{i=1}^{q}\beta_i\sigma_{t-i}^2 \quad (6)$$

where the dummy variable for the negative residuals is defined as

$$S_{t-i} = \begin{Bmatrix} 1 \text{ if } \epsilon_{t-i} < 0; \\ 0 \text{ if } \epsilon_{t-i} \geq 0. \end{Bmatrix} \quad (7)$$

The idea is that asymmetric behaviour of the negative deviations are sources for additional risk. We estimated the AGARCH(1,1) model as follows:

$$\begin{array}{llllll} \hat{\sigma}_t^2 & = 10^{-6}3.3 & + 0.003\epsilon_{t-1}^2 & + 0.18S_{t-1}\epsilon_{t-1}^2 & + 0.86\hat{\sigma}_{t-1}^2. \\ (t-val.) & (4.23) & (0.19) & (7.27) & (44.21) \end{array} \quad (8)$$

The asymmetry parameter γ has a significant t value which is a sign of a rather strong asymmetric persistence effect. The sum of the parameters is less than 1 for positive shocks but is larger than 1 for negative shocks.

2.5 The BEKK model

We want to investigate to what extend multivariate models can improve the variance forecasts. Let $r_t = (r_{1t}, r_{2t})'$ be a vector of returns, then a Gaussian multivariate GARCH model is given by

	AIC	BIC
GARCH(1,0)	-5402	-5374
GARCH(1,1)	-5477*	-5440*
GARCH(2,1)	-5475	-5428
GARCH(2,2)	-5455	-5398

Table 1. AIC and BIC for different GARCH models, MSCI North America index. The star (*) denotes the smallest value.

$$r_t | I_{t-i} \sim N[\mu_t, H_t] \quad , t = 1, ..., T \quad (9)$$

where we assume $\mu_t = \mu$ for the conditional mean and H_t for the conditional covariance matrix given the information set I_t up to time t.

The BEKK(p,q) model of Engle and Kroner (1995) assumes the following parameterisation of the conditional covariance matrix:

$$H_t = A_0A_0' + \sum_{i=1}^{p} A_i(\epsilon_{t-i}\epsilon_{t-i}')A_i' + \sum_{i=1}^{q} B_iH_{t-i}B_i'. \quad (10)$$

where the transposed matrix pairs for each of the coefficient matrices A_i and B_i guarantee symmetry and non-negative-definiteness of the conditional covariance matrix H_t. Using the AIC criteria, we select a BEKK(2,1) model for the returns of the MSCI Europe and MSCI North America indices. For the VaR estimation we use the second diagonal element of the H_t which is the variance of the MSCI North America index. An interesting result is that a bivariate BEKK(1,1) model, where the MSCI World index is used as leading indicator, performs worse with respect to the forecasting performance for the volatility of the MSCI North America index.

3 Forecasting performance

Using the estimation results of the previous two sections and a rolling sample of 800 trading days we forecast the volatility of the returns of the MSCI North America index for the out-of-the-sample period from 25 May 1998 until 3rd April 2000 (we estimate each model 486 times with 800 observations).

We use the auxiliary linear regression approach by Pagan and Schwert (1990) to evaluate the forecasting performance of the volatility models. We simply regress the "realized volatility" (i.e. the squared returns) on a constant and the forecasted volatility:

$$r_t^2 = \alpha + \beta\hat{\sigma_t^2} + \epsilon_t, \quad t = 1, .., T. \quad (11)$$

The constant α should be equal to 0 and the slope β equal to 1. The t-statistic of the coefficients is measure for the bias from the ideal prediction and the R^2 is an overall measure of the forecasting performance. Table 2 summarizes the results of the auxiliary regression (11).

There is a clear sign that multivariate modeling improves the forecasting performance. The BEKK model leads to the smallest "bias" for the intercept and to the largest R^2.

4 VaR comparison

We assume a hypothetical portfolio of 1 Mio US $ which follows the MSCI North America index and the VaR is estimated for the next trading day (in

	α(t-st.)	β(t-st.)	R^2
Naive	0(3.34)	-1.134(-1.47)	0.004
RiskMetrics	0(2.46)	0.575(3.52)	0.025
GARCH(1,1)	0(2.83)	0.571(3.98)	0.032
AGARCH(1,1)	0(3.45)	0.593(5.45)	0.058
BEKK(2,1)	0(0.13)	1.034(5.83)	0.062

Table 2. Model comparison by auxiliary regression of the MSCI North America index (1998/05/25 - 2000/04/03).

the period from 25 May 1998 to 3rd April 2000). Assuming that the returns are normally distributed, the 95%-VaR is computed as 95% quantile of the returns distribution i.e.

$$\widehat{VaR}_{t+1} = -10^6 1.645\hat{\sigma}_{t+1} \tag{12}$$

where $\hat{\sigma}_{t+1}$ is the forecast of the standard deviation given all information up to time t. A 95%-VaR of 1000 US \$ means that with 5% probability the maximum loss in the next trading day will be more than 1000 US \$.

Figure 1 plots the actual portfolio changes ($10^6 r_t$) by tracking the MSCI North America index, the VaR estimates based on the AGARCH(1,1) model and the VaR estimates based on the naive model. The naive model doesn't capture the risk for the first 100 observations and overestimates it in the rest of the period. It is clear that using the VaR estimates of the naive model overestimates the risk if the volatility is small and underestimate it in bear markets.

4.1 Evaluating the VaR by interval forecasts

To evaluate the performance of the different $(1-\alpha)$%-VaR estimators we are using the following four criteria:

1. The failure rate (F)
2. Likelihood ratio test of unconditional coverage (LR_{uc})
3. Likelihood ratio test of independence (LR_{ind})
4. Joint test of coverage and independence (LR_{cc})

We define the failure rate (F) as the number of times for which the actual loss is larger than the estimated VaR. For $t = 1, ..., T$ we define the test statistic

$$F = \sum_{i=1}^{T} D_t \tag{13}$$

with the dummy variable

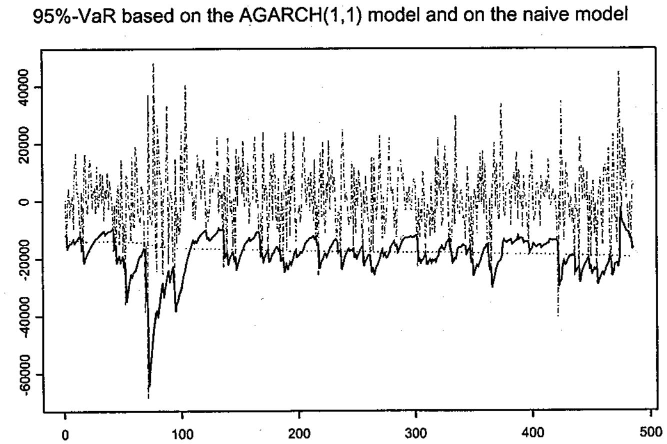

Fig. 1. Actual portfolio changes of the MSCI North America index and the 95%-VaR estimates for the AGARCH(1,1) model and the naive model from 25 May 1998 until 3rd April 2000

$$D_t = \begin{Bmatrix} 1 \text{ if } VaR_t - P_t^a > 0; \\ 0 \text{ if } -VaR_t - P_t^a \leq 0. \end{Bmatrix} \tag{14}$$

The following LR_{uc}, LR_{ind} and LR_{cc} tests are proposed by Christoffersen (1998) to evaluate forecasts over a certain horizon. The likelihood ratio (LR) test of unconditional coverage LR_{uc} tests if $E(D_t) = (1-\alpha)T$ against $E(D_t) \neq (1-\alpha)T$ where α is the probability level for the VaR and T is the number of trading days ($T = 486$) in the evaluation period.

The LR test of independence LR_{ind} tests the hypothesis of independence against a first order Markov chain. Independence would mean that the days for which the actual losses are larger than the estimated value-at-risk ($D_t = 1$) are independent from each other.

The above tests are combined into LR_{cc} test, where the null of the unconditional coverage test is tested against the alternative of the independence tests. The three tests are numerically related by the following equation (see Christoffersen (1998)):

$$LR_{cc} = LR_{uc} + LR_{ind}. \tag{15}$$

The results of the Christoffersen tests for the 90% to 99% VaR based on the BEKK(2,1) model are presented in Figure 2. The three panels plot the LR statistics for the (1-α)-VaR in steps of 1%. The horizontal line corresponds to the 5 per cent critical value of the relevant chi-squared distribution. We see that the null hypothesis is not rejected for α-levels between 3% and 10%. Since the LR_{uc} and LR_{cc} tests are rejected for the 99%-VaR we conclude that the BEKK model might not be we an optimal model for an $\alpha = 1\%$ level.

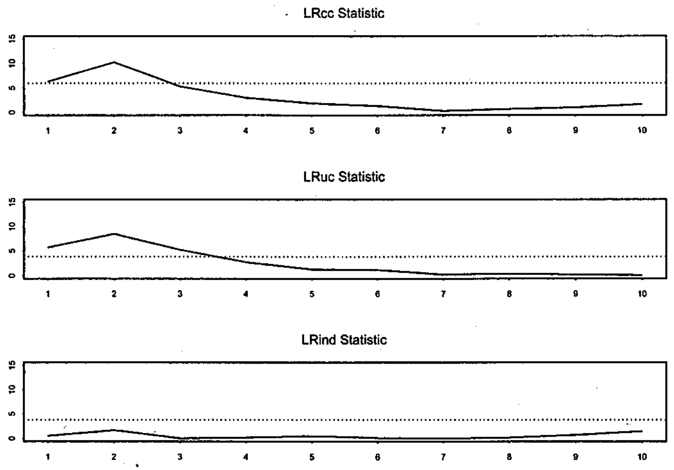

Fig. 2. Christoffersen tests of the $(1-\alpha)$%-VaR (for α between 1% and 10%) for the BEKK(2,1) model of the MSCI North America index

Table 3 summarizes the result for the out-of-the-sample performance of the investigated models for the different indices. The + and - for the Christoffersen tests mean that the null hypothesis is accepted or rejected at a level of 5% respectively. Note that the BEKK(2,1) model leads to the lowest failure rate.

	LR_{uc}	LR_{ind}	LR_{cc}	failures rate (in %)
Naive	-	+	-	39 (8.02)
RiskMetrics	+	+	+	31 (6.38)
GARCH	+	+	+	34 (6.99)
AGARCH	-	+	+	35 (7.20)
BEKK	+	+	+	30 (6.17)

Table 3. 95%-VaR performance for 486 days, MSCI North America index.

5 Conclusions

In this paper we have compared VaR estimates based on volatility forecasts. Therefore we have evaluated the volatility forecasting performance of 3 univariate models together with the historical variance and a bivariate BEKK model. The BEKK model yields the largest R^2, followed by the AGARCH model in the auxiliary regression. Comparing the resulting VaR estimates we found that the multivariate volatility forecasts model yields the best results. The bivariate BEKK model with the MSCI Europe index as a leading indicator leads to the smallest failure rate for 95%-VaR. Surprisingly, the AGARCH model performs worse than the GARCH model with respect to the VaR. The R^2 of the auxiliary regression model can be used as indicator for the goodness of the resulting VaR estimates even if the R^2 values are very small. The Christoffersen tests are useful as diagnostic tests for a good VaR model when the α-level is very small. In a further research paper (see Pojarliev and Polasek (2000)) we extend this approach to evaluate VaR models for stock returns for Europe and the Pacific area.

References

CHRISTOFFERSEN, P. (1990): Evaluationg Interval Forecast. *International Economic Review, 39, 841–862.*

DANIELSON, J. and VRIES, C. (1997): Value-at-Risk and Extreme Returns. In: No 98-017/2 in Tinbergen Institute Discussion Papers from Tinbergen Institute *http://www.tinbergen.nl/discussionpapers/98017.pdf.*

DOCKNER, E. and SCHEICHER, M. (1999): Evaluating Volatility Forecasts and Empirical Distributions in Value at Risk Models. *Finanzmarkt und Portfolio Management, 1, 39–55.*

ENGLE, R. and KRONER, K. (1995): Multivariate Simultaneous GARCH. *Econometric Theory, 11, 122–150.*

MATHSOFT (1996): S-Plus, S+Garch User's Manual: *Data Analysis Products Division, MathSoft, Seattle.*

PAGAN, A. and SCHWERT, W. (1990): Alternative Models for Conditional Stock Volatility. *Journal of Econometrics, 50, 267–290.*

POJARLIEV, M. and POLASEK, W. (2000): Value at Risk estimation for stock indices using the Basle Committee proposal from 1995: *University of Basel, http://www.ihs.ac.at/ polasek.*

RiskMetrics (1996): Risk Metrics Technical Document, fourth edition *http://www.riskmetrics.com/research/.*

ZUCCHINI, W. and NEUMANN, K. (2001): A Comparison of Several Time Series Models for Assessing the Value at Risk of Shares. *Applied Stochastic Models for Business and Industry* 2000, 135–148.