VaR Evaluations Based on Volatility Forecasts of GARCH Models

W. Polasek and M. Pojarliev

Institute of Statistics and Econometrics,
University of Basel, Bernoullistr. 28, CH-4056 Basel, Switzerland
(e-mail: wolfgang.polasek@unibas.ch)

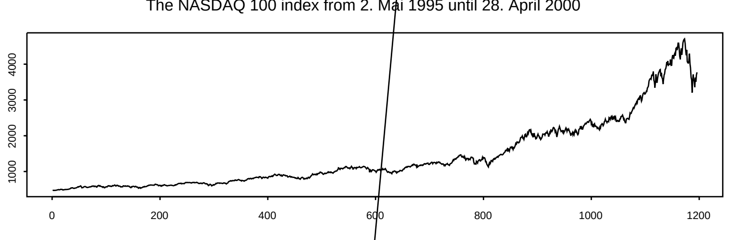
Abstract: In the last years of the 1990'ies many investors were exposed not only to the phenomenon of exceptional or 'abnormal' returns due to stocks of the new economy but also to an increased volatility and risk exposure due to new financial instruments which could affect all international stock markets. The concept of the VaR as the maximum potential loss of a portfolio for the next period for given significance level was established to measure the risk exposures. This paper considers the VaR estimation as a forecasting problem of volatility models and we evaluate the VaR estimates in the same way as we would evaluate volatility forecasts. We focus on VaR models which can be described from classical volatility models and we evaluate them for the returns of the NASDAQ 100 index. We show that the choice of the volatility model can improve the estimation of VaR considerably especially if penalty functions for underestimating the VaR are applied.

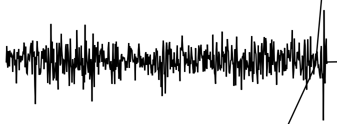
Keywords: GARCH models comparison, NASDAQ 100 index Christofferson test, volatility forecasts, value at risk evaluation.
JEL Classification: C5, G1

1 Introduction

The time of the "dot-shares" boom is over and many investors are now concerned with the risk exposure of their assets. The Yahoo share increased by 2900% in 1995, but in April 2000 the NASDAQ 100 index drop by circa 20%. Even companies with negative earnings per share shot up to 1000% return p.a.. Such volatile behavior of returns increased the demand for new and realistic risk measures for stocks and portfolios. Investors want to know how much is the potential loss, not just how much they can gain. There are numerous proposals and new quantitative approaches to calculate the VaR (see e.g. Jorion (1997)). One of the first commercial approaches has been the variance-covariance approach which was introduced by J.P. Morgan, RiskMetrics (1993). Another group of models are based on simulation techniques. A third group of models applies extreme value theory for VaR estimation (see e.g. Embrechts et al. 1997) with the advantage that a distribution assumption is not required as only the tails are on focus.

The goal of our paper is not to compare different VaR approaches but to investigate the following questions: First, how can time series forecasts improve the VaR estimates and second what types of model leads to better performance. To answer the first question we will compare the performance of different models: 'Naive' models, where the variance estimator is just the historical variance, the RiskMetrics model, GARCH, t-GARCH, an asymmetric GARCH model, the power GARCH and exponential GARCH model. In second step we will evaluate the resulting VaR estimators by the Christofferson (1998) test. We also calculate the costs of the VaR requirements for banks with and without a penalty function. Naive models have lower VaR and require smaller potential capital costs if no penalty function for failures of meeting the target risk is imposed. We show that imposing simple penalty functions will make GARCH based VaR evaluations more desirable than naive models.
The paper is organized as follows: In the next section we estimate the volatility of the NASDAQ 100 index with the different models and compare their forecasting performance. In section 3 we estimate the VaR of hypothetical portfolio of 1 Mio. $ which is invested in QQQ shares which tracks the NASDAQ 100 index. In the last section we conclude.





2 Volatility forecasts

We will investigate the volatility of the daily returns of the NASDAQ 100 index from 2nd May 1995 until 28 April 2000. The first 800 observations (from 2nd May 1995 until 25 September 1998) are used for the model selection an the rest for out-of- sample comparison. Figure 1 plots the labels and the returns for the whole time horizon.

2.1 Volatility models

In this section we describe all naive and GARCH based volatilty model which we will use for the evaluation of the VaR in the next section.

Naive model

We are using the variance of a moving sample of 800 observation (approximately 3 years) as forecast for the next period (1000 observations were suggested by Dockner and Scheicher, 1999). The volatility forecast for time $t+1$ is based on the (historical) sampling variance for a time length of N observations (e.g. 800 days) and is computed by

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

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

The RiskMetrics model

The model proposed by J.P. Morgan (1995) is an exponentially weighted moving average model. The volatility of the next period is calculated as a MA process of weighted squared deviations from the mean, where the weights decay exponentially:

$$\sigma^2_{t+1} = (1-\lambda) \sum_{i=1}^{\infty} (r_{t+1-i} - \mu)^2 = \lambda \sigma^2_t + (1-\lambda)(r_t - \mu)^2 \tag{2}$$

where λ is the weight factor. As proposed by RiskMetrics we set λ equal to 0.94. As an initial value for σ^2_t, we use the sample variance (for 800 trading days) and for μ we use $\hat{\mu}$.

The GARCH(p,q) model

Using the AIC and BIC criteria we select a GARCH(1,1) model for the in-sample period from 2nd May 1995 until 25 September 1998. The estimated coefficients of the variance equation are given below (with t-statistic are given in parentheses):

$$\begin{array}{ccccccc} \hat{\sigma}^2 & = & 0.00002161 & + & 0.0845\hat{\epsilon}^2_{t-1} & + & 0.84023\hat{\sigma^2}_{t-1} \\ & & (2.657) & & (3.935) & & (22.71) \end{array} \quad (3)$$

Table 1 shows the AIC and BIC for different type of models.

	AIC	**BIC**
GARCH(1,0)	-4269	-4241
GARCH(0,1)	-4264	-4234
$GARCH(1,1)^*$	-4291	-4254
GARCH(2,1)	-4289	-4242
GARCH(2,2)	-4282	-4226
AR(1)-GARCH(1,1)	-4278	-4212

Table 1: GARCH model selection: AIC and BIC for different GARCH models. (The * indicates the best model.)

The conditional distribution of ϵ_t is assumed to be Gaussian normal with mean zero and conditional variance σ_t^2. The autocorrelation function (ACF) of the squared returns are shown in Figure 2. We can observe substantial correlation which motivate the use of GARCH modeling. The ACF of the squared standardized residuals doesn't show such correlation structure which shows that the model diagnostics for the GARCH (1,1) model is satisfactory.

GARCH model with t-distribution

Previous research has shown that the conditional distribution for the error term in return models often has heavier tails than the Gaussian distribution. Therefore we estimate a t-GARCH(1,1) model where the t distribution is used as conditional distribution for the residuals in the mean equation of the t-GARCH model. The estimated variance equation of the t-GARCH model is given by

$$\begin{array}{ccccccc} \hat{\sigma}_t^2 & = & 0.000019 & + & 0.0679\epsilon^2_{t-1} & + & 0.8360\sigma^2_{t-1}. \\ (t-st.) & & (1.87) & & (2.84) & & (13.79) \end{array} \quad (4)$$

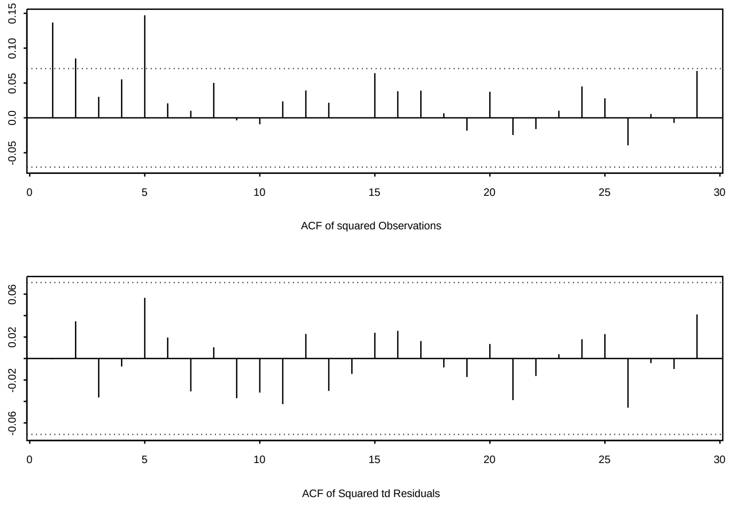


Figure 2: Model diagnostics: The ACF of the squared returns (upper panel) and squared standardized residuals of the GARCH(1,1) model.

The asymmetric GARCH model

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

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} (\alpha_i + \gamma_i S_{t-i}) \epsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2 \tag{5}$$

where S_{t-i} is the dummy variable for the negative residuals and is defined as

$$S_{t-i} = \begin{array}{lll} 1 & \text{if} & \epsilon_{t-i} < 0; \\ 0 & \text{if} & \epsilon_{t-i} \geq 0. \end{array}$$

The idea is that an asymmetric risk behaviour can be present in the variance equation and the negative residuals are sources for additional risk. We estimate the AGARCH(1,1) model with the variance equation given by

$$\begin{array}{ccccccccc} \hat{\sigma_t^2} & = & 10^{-5}3 & - & 0.017\epsilon_{t-1}^2 & + & 0.174 S_{t-1}\epsilon_{t-1}^2 & + & 0.817\hat{\sigma}_{t-1}^2 \\ (t-st.) & & (3.29) & & (-0.53) & & (4.05) & & (18.75) \end{array} \tag{6}$$

The exponential GARCH model

An exponential GARCH model (EGARCH) uses a Gaussian distribution for the errors of the mean equation and has the following specification for the variance equation

$$\ln \hat{\sigma}_t^2 = a + \sum_{i=1}^{p} \alpha_i \ln \hat{\epsilon}_{t-i}^2 + \sum_{i=1}^{q} \beta_i \ln \hat{\sigma}_{t-i}^2 \tag{7}$$

where $\hat{\epsilon}_t = y_t - \hat{\mu}_t$ is the estimated residual of the mean equation. The variance equation for the in-sample-period was estimated as

$$\begin{array}{llllllll} \widehat{\ln \sigma^2} & = & -1.1922 & + & 0.1881 \ln \epsilon_{t-1}^2 & + & 0.8728 \ln \sigma_{t-1}^2 \\ (t-st.) & & (-2.92) & & (4.64) & & (17.93) \end{array} \tag{8}$$

The power GARCH model

The power GARCH model (PGARCH) with common exponent d is defined in analogy to the Box-Cox transformation as

$$\sigma_t^d = a + \sum_{i=1}^{p} \alpha_i \mid \epsilon_{t-i} \mid^d + \sum_{i=1}^{q} \beta_i \ln \sigma_{t-i}^d \tag{9}$$

where the exponent d can be estimated or specified apriori. We have estimated for the in-sample-period the exponent by $\hat{d}$ over a grid $d \in (-2, 2)$ and we found

$$\begin{array}{llllllll} \hat{\sigma}^{\hat{d}} & = & 0.000013 & + & 0.0788 \mid \epsilon_{t-1} \mid^{\hat{d}} & + & 0.8433 \sigma_{t-1}^{\hat{d}} \\ (t-st.) & & (0.433) & & (3.09) & & (21.36) \end{array} \tag{10}$$

with $\hat{d} = 2.117$ and a t-value of (3.61). Note that the exponent is defined since all variables including the absolute values of the residuals are positive and the estimate is close to the 'natural' power 2 which indicates that the variance is the driving risk factor in power GARCH models for NASDAQ returns.

3 VaR comparison

We assume a hypothetical portfolio of 1 Mio. US \$ invested in QQQ, a share which tracks the NASDAQ 100 index and the VaR is estimated for each trading day. As we don't investigate the role of the distributional assumption for daily returns, we focus on the role of volatility predictions

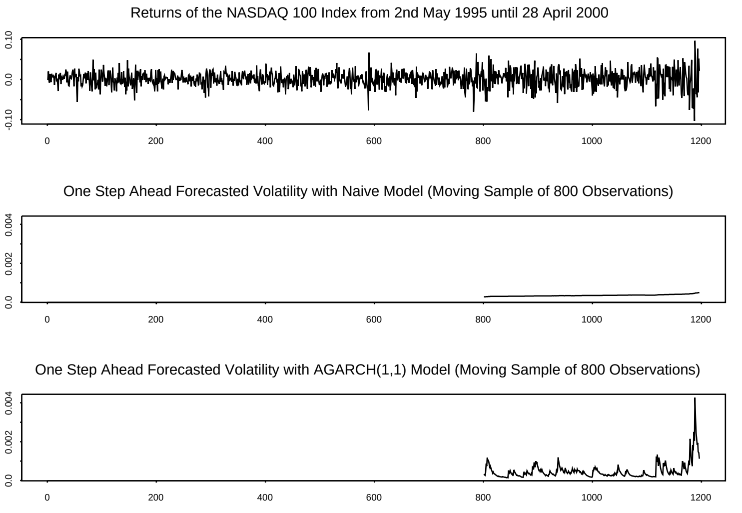


Figure 3: Naive and GARCH models: Volatility forecasts for the NASDAQ 100 index for out-of-the-sample period from 26 September 1998 until 28 April 2000

on the VaR. Therefore we assume that the returns of the NASDAQ 100 index are normally distributed and we estimate as 95% VaR the 5% quantile of the normal density. For the portfolio of size A, e.g. $A = 1$ Mio. US$) the VaR is computed as multiple of the volatility forecast $\hat{\sigma}_{t+1}$ for the next period:

$$\widehat{VaR}_{t+1} = At(\alpha)\hat{\sigma}_{t+1}. \tag{11}$$

where $t(\alpha)$ is the α-quantile of the normal density. For $\alpha = 0.05$ the normal quantile is $t(\alpha) = -1.65$. Figure 4 plots the actual portfolio changes (daily returns: $P_t^a = 10^6 r_t$) and the VaR estimates for GARCH(1,1) model and the 'Naive' model by tracking the NASDAQ 100 index returns.

3.1 Evaluation criteria

To evaluate the performance of the different VaR estimators, we are using the following five criteria:

1. Forecasting performance of the VaR

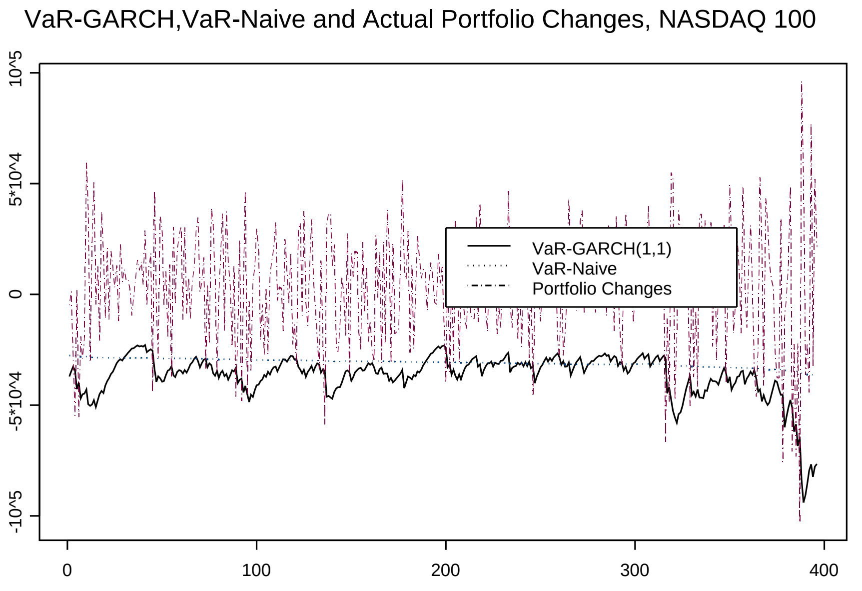


Figure 4: 95% VaR performance: Actual portfolio changes for the NASDAQ 100 index, the VaR-GARCH and VaR-Naive estimates from 25 May 1998 until 28 April 2000.

2. The failures F and the failure rate ($F\%$)
3. Likelihood ratio test of unconditional coverage (LR_{uc})
4. Likelihood ratio test of independence (LR_{ind})
5. Joint test of coverage and independence (LR_{cc})
6. Average and penalized VaR "costs"

The following sections will discuss these criteria for the example of the NASDAQ 100 index.

3.2 Forecasting performance of the VaR

Using the variance equations of the GARCH family of the last section, i.e. (1), (2), (6), (4), (8) and (10) and a moving sample of 800 trading days to forecast the volatility of the returns of the NASDAQ 100 index for out-of-the-sample period from 26 September 1998 until 28 April 2000 (396 trading days). Figure 3 shows the result for the naive model and the AGARCH model. We see that the volatility forecasts are numerically and qualitatively quite different but in the next section we want to find out if this has a significant effect on the VaR performance. Following Dockner and Scheicher (1999) we are using the linear regression approach by Pagan and Schwert (1990) to evaluate the forecasting performance of the volatility models. We simply regress the "realized volatility" (proxied by squared returns) on a constant and the forecasted volatility for a certain comparison period $t = 1, ..., n$,

$$r_t^2 = \alpha + \beta \hat{\sigma_t^2} + \epsilon_t \tag{12}$$

In this auxiliary regression the constant should be close to 0 and the slope close to 1. The t-statistic of the coefficients is a measure for the bias and the squared correlation (R^2 or R squared) is measure of the forecasting performance. Table 2 summarizes the result of the auxiliary regression (12) for different volatility models.

Table 2 exhibits that the 'naive' model performs much worse than the time series models. According to the R^2 criterion the AGARCH model dominates the other volatility models.

	α	(t-st.)	β	(t-st.)	R^2
Naive	0.002	(-4.82)	8.24	(6.31)	0.09
RiskMetrics	-0.001	(-4.23)	4.41	(7.37)	0.12
GARCH(1,1)	0	(0.40)	1.19	(8.95)	0.17
t-GARCH(1,1)	0	(0.78)	1.28	(8.96)	0.17
AGARCH(1,1)	0	(0.21)	1.33	(11.79)	0.26
EGARCH(1,1)	0	(-0.95)	1.58	(8.59)	0.16
PGARCH(1,1), $p = 2.12$	0	(0.27)	1.24	(8.05)	0.14

Table 2: Forecasting performance for volatilities: Summarizing the auxiliary regression (12)

3.3 Failures and failure rates

The failures (F) are defined as the number of time for which the actual portfolio return is smaller than the estimated VaR. For $t = 1, ..., T$ and $T = 394$ the failures are calculated as

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

and the failure rate ($F\%$) is given as $F = F/T$. The dummy variables are defined as

$$D_t = \begin{cases} 1 & \text{if} \quad P_t^a < VaR_t; \\ 0 & \text{else} \end{cases} \tag{14}$$

3.4 Unbiased and independent VaR forecasts

The LR_{uc}, LR_{ind} and LR_{cc} likelihood ratio (LR) tests are proposed by Christoffersen (1998) for general interval forecasts. The LR test of unconditional coverage can be used for GARCH based VaR models to test if the VaR failures are unbiased, i.e. $E(D_t) = \alpha T$ against $E(D_t) \neq \alpha T$, where α is the significance level for the VaR and T is the number of trading days in the evaluation period.
The LR test of independence tests the hypothesis of independence against a first order Markov chain. Independence would mean that the days for which the actual losses are larger in absolute value than the estimated value-at-risk (then the dummy variable in (14) is set $D_t = 1$) are independent from each other.
The above tests are combined into the LR_{cc} test where the null hypothesis of conditional coverage is tested with the LR_{cc} statistic. The LR_{cc} test can be computed as sum of the previous two tests (see Christoffersen, 1998):

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

How good are the Christoffersen tests for VaR forecasts?

The results of the Christoffersen likelihood ratio tests for the VaR forecasts of the GARCH(1,1) model for the evaluation period from 25 May 1998 until 28 April 2000 are presented in Figure 5. The three lines in each panel connect the values of the LR statistics from the VaR-1% to the VaR-10% in 1% steps. The horizontal solid line corresponds to the 5 percent critical value of the relevant chi-squared distribution.
We see that the independence test is met for all 10 percentage steps of the VaR evaluation. The LR test of unconditional coverage does not reject the null hypothesis in the range of the VaR between 2% and 5%. Therefore the conditional coverage test is also only good (i.e. not rejected) in this interval from 2% to 5%, since LR_{cc} in (15) is defined as the sum of the unconditional and the independence test.
There are two possible reasons to explain this phenomenon: First, the Christoffersen test is not applicable for VaR evaluation which is rather unlikely since the test on unbiasedness and independence of forecasts

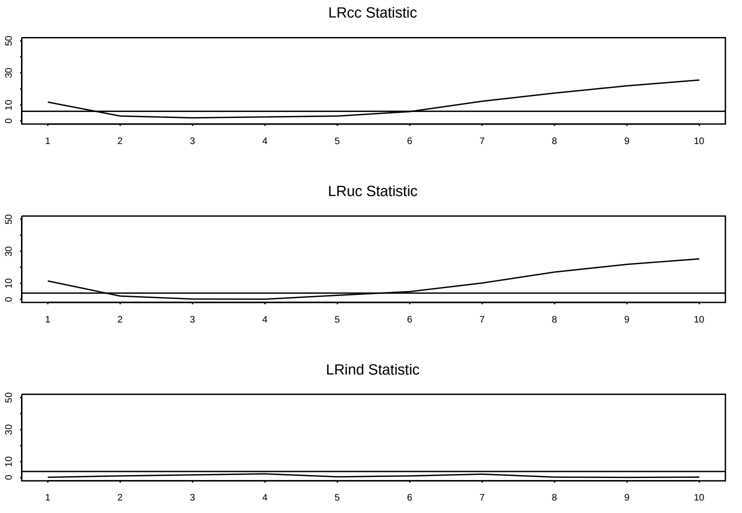


Figure 5: VaR-1% to VaR-10% performance using the likelihood ratio statistics of Christoffersen (1998)

seems rather plausible. Second, GARCH based VaR models are only reliable for special situations. GARCH models are good to capture volatility clusters but they need the triggering event for creating a volatility cluster. It seems that for high and low end VaR the volatility forecasts are too optimistic and are surprised by the force of a new volatility shock. It remains to be seen if other classes of volatility models experience similar evaluation patterns.

Table 3 summarizes the results for the out-of-the-sample 5% VaR performance based on volatility forecasts of GARCH models.

3.5 VaR costs and penalties

The regulatory requirement for a 99% VaR is based on a 10-day average. Therefore we suggest to use the average sum of VaR in % of the portfolio as an average "VaR costs" criterion. A model which lead to lower costs will be desirable for banks but this has to be traded off with the penalty which has to be paid if the VaR is not met over a certain time period. On the other side it is clear that an overestimation of VaR will lead to an increase in the opportunity costs. This gives simple models like the naive or the RiskMetrics model an advantage. But the penalty function has to be of right amount and a monotone function of the failure rate

to beat the naive models.
From Table 3 we see that the GARCH(1,1) model produces the best VaR estimates in terms of the failure rates. With 6.81 % it is still above the target of 5% but the naive model is with 10.5% more than double the size off the target level. Not surprisingly, the naive models have the lowest average VaR costs. But introducing a penalty factor will offset this advantage easily. Most penalty function are based on penalty factors which depend on the number of overshooting the VaR during a certain time period and are based on a 99% VaR.
We suggest a simple penalty function which increases the average VaR costs by 5% (or the factor .05) for each failure above the target of 5 failures in the evaluation period. Note that penalty factor increases in the same step size as it is described in Zucchini and Neuman (2001) for the German regulations. (The step change has to be substantial, else it would be not effective and would be dominated by the average costs.) This leads to the formula

$$VaR_{pc} = costs * max(1, 1 + \frac{failures - 5}{20})$$

Also we see from the last column of Table 3 (i.e. the VaR_{pc} = VaR with penalty costs) that the VaR_{pc} with penalty is the lowest for the GARCH(1,1) model, the model with the lowest number of failures.

4 Conclusions

The VaR evaluation study based on GARCH forecasts shows that even with the simple assumption that daily returns are normally distributed, the use of conditional variance forecasts can improve the VaR estimates considerably. The advantage becomes clear if the costs for the VaR is penalized and the failure rate of the VaR is high. Other proposals for

Model	LR_{uc}	LR_{ind}	LR_{cc}	failures F (and F%)	"VaR " costs	VaR$_{pc}$
Naive	-	+	-	43(10.8)	3.08*	8.93
RiskMetrics	-	+	-	42(10.6)	3.15	8.98
GARCH	+	+	+	27(6.81)*	3.66	7.69*
TGARCH	+	+	+	28(10.8)	3.62	7.78
AGARCH	-	+	+	30(7.57)	3.45	7.76
EGARCH	-	+	+	30(7.57)	3.53	7.94
PGARCH	+	+	+	28(7.07)	3.63	7.80

Table 3: 5%-VaR performance for the NASDAQ 100 returns based on volatility forecasts. The star * marks the best model.

the penalty function when the VaR requirements for banks are not met can be found in Pojarliev and Polasek (2000). Using the Christofferson test to evaluate the VaR forecasts we see that for 1% VaR the hypothesis of unbiased coverage is violated. This might be an indication that GARCH based VaR evaluations have to be improved for high and extreme coverage probabilities to obtain improved VaR estimates. Further research will show if outlier models (see e.g. Polasek and Jin (1997) or other distributional assumption have an effect on VaR performances and how in general volatility based VaR approaches compare with extreme value based VaR approaches.

References

Basle Committee on Banking Supervison, (1996): Amendment on the capital accord to incorporate market risks.
http://www.bis.org./publ/index.htm

Bollerslev, T.(1986): Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31, 307-327.*

Christoffersen, P. (1998): Evaluating Interval Forecast.
http://www.ssc.upenn.edu/peterfc/

Dockner, E. and Scheicher, M. (1999): Evaluating Volatility Forecasts and Empirical Distributions in Value at Risk Models. *Finanzmarkt und Portfolio Management, 13, Nr.1, pp. 39-55.*

Embrechts P., Klüppelberg C., and Mikosch T. (1997) Modelling Extremal Events for Insurance and Finance, Springer Verlag, Berlin.

Mathsoft, Inc. (1996): S+GARCH User's Manual. Data Analysis Products Division, MathSoft, Seattle.

Morgan J.P. (1996): Risk Metrics - Technical Document, 4th edition.
http://www.riskmetrics.com/rm/pubs/index.cgi

Pagan, A. and Schwert, W. (1990): Alternative Models for Conditional Stock Volatility. *Journal of Econometrics 50, pp. 267-290.*

Jorion P.(1997) Value at Risk, MxGraw-Hill, NY.

Pojarliev M. and Polasek W. (2000) Value at Risk Estimation for Stock Indices Using the Basle Committee Proposal from 1995, WWZ Basel, mimeo.

Polasek W. and Jin S. (1997). GARCH Models with Outliers. In: *Classification and Knowledge Organisation,* Klar R. and Opitz O. (eds.) Springer Verlag, 178-186.

W.andagan,)