Keskustelualoitteita #52
Joensuun yliopisto, Taloustieteet

GARCH Modelling with Power Exponential Distribution - Applications to Value at Risk Estimation

Mika Louhelainen
&
Jukka Nyblom

ISBN 978-952-219-072-7
ISSN 1795-7885
no 52

GARCH Modelling with Power Exponential Distribution—Applications to Value at Risk Estimation

Mika Louhelainen *
University of Joensuu

Jukka Nyblom †
University of Jyväskylä

November 2007

Abstract

We investigate the performance of the GARCH modelling strategy with symmetric and asymmetric power exponential error distributions in predicting VaR values. Some elegance of formulation is gained by expressing the volatility recursion in terms of the power characterizing the power exponential error distribution. At the same time useful asymptotic results become readily available. Our approach is applied to eight series of daily returns of lengths around 2800. Our overall conclusion is that many types of GARCH models capture the volatility dynamics adequately. Nevertheless, more reasonable estimates for actual VaR values are obtained with bootstrap than with the estimated error distribution.

Key words: bootstrap, empirical finance, fat tails, volatility modelling.
JEL classification: C22, C51, C53.

*Department of Economics and Business Administration, PO Box 111, FI-80101 Joensuu, Finland, email: mika.louhelainen@joensuu.fi

†Department of Mathematics and Statistics, PO Box 35 (MaD), FIN-40014 Jyväskylä, Finland, email: jukka.nyblom@jyu.fi

1 Introduction

Over the last decade Value at Risk (VaR) has become most important instrument to measure the market risk of the financial institutions. One reason is that the Basel II, which directs banking laws and regulations world wide, recommends that a capital adequacy of institutions should be based VaR modelling. The VaR measure is easy to work with, which is the other reason for the popularity. In fact, Value at Risk is an attempt to provide a single number summarizing the total market risk in an asset or even investment portfolio of the financial institution. Jorion (1997) has defined VaR exposure as follows:" VaR is the worst expected loss over a great horizon within a given confidence level". Statistically VaR is a quantile of the expected return distribution of the asset. In order to be specific choose the quantile, $q(0.99)$ say, corresponding to the probability 0.99, and let us consider the possible loss occurring on the next day. Then with probability 0.99 the next day's loss will be no greater than $q(0.99)$. VaR can be determined both long and short positions. If a trader is holding a long (short) position she is interested in the left (right) tail of the distribution.

Therefore, in VaR calculation we are interested in the tail behavior of the expected asset return distribution. The financial series typically display high kurtosis, fat tails and negative skewness, and most importantly they exhibit clustering volatility. Granger and Ding (1995) listed a few more such features. The most easiest way to estimate the return distribution is the use of the past return data in a very direct way as a guide to what may happen in the future. This historical simulation approach has the advantage that we do not have to make an assumption on the return distribution. But there are disadvantage which Hull and White (1998) noticed. Historical simulation does not easily allow volatility updating schemes to be used. In practice we would also need a large database for historical simulation.

The parametric approaches such as exponentially weighted moving averages (EWMA) and generalized autoregressive conditionally heteroscedastic (GARCH) models are probably the most common tools to determine VaR for linear assets such as bonds and stocks. The main disadvantage of these approaches is that we have to make an assumption on the error distribution. Commonly a normal distribution is used regardless of apparent conflict with the data. The consequence of this approach is that VaR is underestimated due to the short tails of the normal distribution. Instead of the normal distribution, other distributions are also used, e.g. see Giot and Laurent (2004), Kuester, Mittnik and Paolella (2006) and Komunjer (2007). Our choice for the error distribution is the same as that of Komunjer (2007), i.e. the power exponential distribution and its asymmetric version. The difference is that she focuses on expected shortfall whereas we consider Value at Risk. Another deviance is that our volatility dynamics is modelled in terms of the conditional expectation of the λ-th moment with λ being the exponent in the error distribution. Komunjer (2007) uses conditional variances. Our approach yields, in addition to simplified formulas, certain stationary results as well as asymptotic results.

We propose to use robust rank correlations to protect ourselves against outliers in checking residual autocorrelations. We also strongly encourage to use graphical techniques (QQplots and symmetry plots) rather than formal statistical tests in assessing the adequacy of the error distribution.

The remainder of the paper is organized as follows: Section 2 introduce the asymmetric power exponential GARCH model and Section 3 describes the maximum likelihood estimation of this model. In Section 4 concentrates on VaR application and results from the empirical investigation. In Section 5 empirical results are discussed.

2 Asymmetric power exponential model

The standard GARCH(p, q) model introduced by Bollerslev (1986) is defined by the equations

$$
\begin{aligned}
y_t &= \sigma_t \varepsilon_t, && (1)\\
\sigma_t^2 &= \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, && (2)\\
t &= 1, \dots, n,
\end{aligned}
$$

where the constants satisfy the nonnegative constraints $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$ for $i = 1, \dots, p$, $j = 1, \dots, q$. A standard assumption for the errors ε_t is that they are independent and identically distributed (i.i.d.) with mean zero and unit variance. Assuming unit variance means no loss of generality, but if $E(\varepsilon_t) = m \neq 0$, then the conditional expectation of y_t given the past values y_{t-j}, $j = 1, 2, \dots$, is $m\sigma_t$ yielding a type of risk premium with parameter m. Under mild conditions Bougerol and Picard (1992a and 1992b) have shown that recursions (1) and (2) define a unique strictly stationary process. The conditions allow $E(\varepsilon_t) \neq 0$. A simple sufficient condition for stationarity is that $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$ already established by Bollerslev (1986). When all α_i and β_j are positive also the IGARCH process with $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j = 1$ is stationary.

Ding, Granger and Engle (1993) introduced a model where the recursion (2) occurs not with power two but with some other power λ, and where negative and positive errors at time t affect differently on the future volatilities. The model, denoted here by APEGARCH(p, q) is

$$y_t = \sigma_t \varepsilon_t, \tag{3}$$

$$\sigma_t^\lambda = \alpha_0 + \sum \alpha_i |y_{t-i} - \gamma |y_{t-i}||^\lambda + \sum \beta_j \sigma_{t-j}^\lambda. \tag{4}$$

Ding et al. (1993) call it APARCH under the assumption that the errors ε_t are independent unit normal variables. Commonly we have restrictions $1 \leq \lambda \leq 2$. We use the acronyms PEGARCH when $\gamma = 0$, AGARCH when $\lambda = 2$ and GARCH when $\gamma = 0$ and $\lambda = 2$.

Following He and Teräsvirta (1999), *mutatis mutandi*, we can make the transformations $u_t = \text{sign}(y_t - \gamma|y_t|)|y_t - \gamma|y_t||^{\lambda/2}$ and then define $\eta_t = \text{sign}(\varepsilon_t - \gamma|\varepsilon_t|)|\varepsilon_t - \gamma|\varepsilon_t||^{\lambda/2}$ with $\tau_t = \sigma_t^{\lambda/2}$. Formally, the standard GARCH model appears with observations u_t driven by new errors η_t and with volatilities τ_t^2. If the conditions of Bougerol and Picard (1992a and 1992b) for the stationary solution exist in the transformed model with u_t, τ_t^2 and η_t, then automatically the original model, with y_t, σ_t^λ and ε_t, admits a unique strictly stationary solution. In order to estimate the parameters in (3) and (4) we need an assumption for the distribution of the errors. Commonly a normal distribution is taken for this purpose. But in applications we most often are faced with the facts that the error distribution has thicker tails than the normal distribution, and in addition it may exhibit skewness. Yet, an adequate description of the dynamics of the volatility process σ_t^2 may be quite satisfying also under the tentative, plausibly wrong, normal assumption. But, e.g. in VaR estimation, the distributional assumptions are more crucial.

Our choice for the distribution of ε_t is the asymmetric power exponential distribution APE$(0, 1, \lambda, \gamma)$ given by the density

$$f(x; \lambda) = \frac{1 - \gamma^2}{2\Gamma(1 + 1/\lambda)\lambda^{1/\lambda}} \exp\left(-\frac{|x - \gamma|x||^\lambda}{\lambda}\right), \ -\infty < x < \infty, \ \lambda > 0, \quad (5)$$

which nicely matches with (3) and (4). This distribution is also called an asymmetric generalized error distribution. In literature it is parameterized in various ways (e.g. see Komunjer, 2007, references therein), but our specification (5) has some benefits as seen later.

We find immediately that the symmetric versions with $\gamma = 0$ lead to the standard normal density when $\lambda = 2$, and to the Laplace distribution when $\lambda = 1$. Moreover, in Appendix we will see that if $\varepsilon_t \sim$ APE$(0, 1, \lambda, \gamma)$, then $E(|\varepsilon_t - \gamma|\varepsilon_t||^\lambda) = 1$ leading to the conditional expectation $E(|y_t - \gamma|y_t||^\lambda \,|\, \mathcal{F}_{t-1}) = \sigma_t^\lambda$, where $\mathcal{F}_{t-1}$ is the σ field induced by the past values $y_{t-1}, y_{t-2}, \ldots$.

3 Maximum likelihood

In practice the model often has also a mean process μ_t (measurable with respect to $\mathcal{F}_{t-1}$). Then the model (3) and (4) applies to the difference $y_t - \mu_t$. In our applications we assume a constant mean μ.

For simplicity of notation write $r_t = y_t - \mu - \gamma|y_t - \mu|$. The log-likelihood function corresponding to APEGARCH model is then

$$\log L = n \log C(\gamma, \lambda) - \frac{1}{\lambda} \sum \left(\log \sigma_t^\lambda + \frac{|r_t|^\lambda}{\sigma_t^\lambda} \right). \quad (6)$$

where $C(\gamma, \lambda)$ is the scaling factor in (5). The MLE is found as a solution to

$$\frac{\partial \log L}{\partial \phi} = 0, \quad (7)$$

where ϕ comprises all the parameters which need to be estimated. Appendix A.1 provides the formulas for the partial derivatives for APEGARCH(1,1).

The asymptotic theory for the APEGARCH models has not been strictly proved, but a partial results are obtainable from Berkes, Horváth and Kokoszka (2003). Their results concern with ordinary GARCH with error distribution satisfying mild regularity conditions. Suppose that λ, μ and γ are known. The He-Teräsvirta transformation $u_t = \text{sign}(r_t)|r_t|^{\lambda/2}$ makes the likelihood equal to the Gaussian likelihood (apart from constant), and we can deduce that under analogous regularity conditions the quasi-maximum likelihood estimates for the parameters $\theta = (\alpha_0, \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q)'$, i.e. solution from (9) with λ, μ and γ known, are asymptotically multivariate normal. Let this quasi-MLE be $\widehat{\theta}_n$. Then we have

$$
\begin{aligned}
\sqrt{n}(\widehat{\theta}_n - \theta) &\xrightarrow{\mathcal{D}} N(0, v^2 B(\theta)^{-1}), \\
B(\theta) &= E\left[\frac{1}{\sigma_t^{2\lambda}} \frac{\partial \sigma_t^\lambda}{\partial \theta} \frac{\partial \sigma_t^\lambda}{\partial \theta'}\right], \\
v^2 &= \text{var}(|\varepsilon_t|^\lambda).
\end{aligned}
$$

The expectation is with respect to the stationary distribution of the process. An easy calculation shows that $v^2 = \lambda$ leading to the well known result $v^2 = 2$ under normal errors.

In practice $B(\theta)$ is estimated by the corresponding sample average at $\theta = \widehat{\theta}_n$. Also, instead of $v^2 = \lambda$ we can use the sample variance of the absolute residuals raised to the power λ.

Despite of the lack of a rigorous proof we have used the standard maximum likelihood theory in our applications in other words we have assumed that in large samples

$$
\widehat{\phi} \sim N(\phi, n^{-1}\Omega),
$$

with $\widehat{\phi}$ being the MLE and Ω as the limit of

$$-\left(\frac{1}{n}\frac{\partial^2 \log L(\phi)}{\partial\phi\partial\phi'}\right)^{-1}.$$

An alternative expression for Ω is obtained from the quasi maximum likelihood theory and is the limit of

$$\left(\frac{1}{n}\frac{\partial^2 \log L(\phi)}{\partial\phi\partial\phi'}\right)^{-1}\left(\frac{1}{n}\sum\frac{\partial \log \ell_t(\phi)}{\partial\phi}\frac{\partial \log \ell_t(\phi)}{\partial\phi'}\right)\left(\frac{1}{n}\frac{\partial^2 \log L(\phi)}{\partial\phi\partial\phi'}\right)^{-1},$$

where ℓ_t is the conditional log-likelihood of y_t, i.e. $\log L = \sum \ell_t$. In practice we replace ϕ by the estimate $\widehat{\phi}$ when computing standard errors.

4 Applications

4.1 Value at Risk

Value at Risk (VaR) is mainly concerned with market risk which is one type of risk in financial markets. VaR has been increasingly used as a risk management tool (see e.g., Jorion, 1997, and Tsay, 2007, Ch 7).

Suppose that at the time t we are interested the risk of the financial position for the next k periods. Let the price of the financial position be P_t at time t. Then the change in value of our position is $P_{t+k} - P_t$, and the VaR of a long position is defined to be $\mathrm{VaR}(t,k,p)$ satisfying

$$\mathbb{P}[P_{t+k} - P_t \leq \mathrm{VaR}(t,k,p) \mid \mathcal{F}_t] = p. \tag{8}$$

Typically $\mathrm{VaR} = \mathrm{VaR}(t,k,p)$ is negative for small p. Therefore (8) defines the probability p that the holder of a financial asset suffers a loss which is greater than or equal to $-\mathrm{VaR}$ (taken positive now). Alternatively, with

Table 1: Parameter estimates of APEGARCH and PEGARCH models

Series	Method	α_0	α_1	β_1	λ	μ	γ
Apple	APEGARCH	0.012	0.018	0.978	1.211	0.000	0.026
		(0.009)	(0.006)	(0.008)	(0.040)	(0.013)	(0.013)
	PEGARCH	0.011	0.018	0.978	1.207	0.055	0
		(0.008)	(0.006)	(0.008)	(0.039)	(0.057)	(-)
Barclays	APEGARCH	0.014	0.050	0.943	1.303	0.018	0.018
		(0.006)	(0.008)	(0.010)	(0.050)	(0.042)	(0.019)
	PEGARCH	0.014	0.050	0.943	1.309	0.052	0
		(0.006)	(0.008)	(0.010)	(0.048)	(0.030)	(-)
British Airways	APEGARCH	0.023	0.062	0.930	1.392	-0.005	0.018
		(0.010)	(0.011)	(0.013)	(0.046)	(0.044)	(0.018)
	PEGARCH	0.023	(0.061)	(0.930	(1.392	0.023	0
		(0.010)	(0.011)	(0.013)	(0.046)	(0.034)	(-)
Dow Jones	APEGARCH	0.009	0.070	0.921	1.529	0.116	-0.047
		(0.003)	(0.010)	(0.011)	(0.057)	(0.028)	(0.021)
	PEGARCH	0.009	0.069	0.921	1.515	0.063	0
		(0.003)	(0.010)	(0.011)	(0.055)	(0.016)	(-)
Microsoft	APEGARCH	0.012	0.063	0.932	1.306	-0.106	0.063
		(0.006)	(0.011)	(0.012)	(0.042)	(0.026)	(0.015)
	PEGARCH	0.012	0.060	0.935	1.310	0.009	0
		(0.005)	(0.011)	(0.012)	(0.042)	(0.031)	(-)
NASDAQ	APEGARCH	0.011	0.088	0.908	1.663	0.225	-0.078
		(0.004)	(0.012)	(0.012)	(0.062)	(0.034)	(0.019)
	PEGARCH	0.012	0.088	0.908	1.678	0.107	0
		(0.004)	(0.011)	(0.012)	(0.062)	(0.021)	(-)
Nokia	APEGARCH	0.012	0.042	0.954	1.270	0.231	-0.031
		(0.007)	(0.008)	(0.010)	(0.038)	(0.096)	(0.025)
	PEGARCH	0.012	0.043	0.954	1.265	0.141	0
		(0.007)	(0.008)	(0.010)	(0.038)	(0.030)	(-)
Shell	APEGARCH	0.010	0.052	0.942	1.373	0.140	-0.032
		(0.005)	(0.010)	(0.011)	(0.047)	(0.038)	(0.020)
	PEGARCH	0.010	0.052	0.942	1.378	0.088	0
		(0.005)	(0.010)	(0.011)	(0.047)	(0.025)	(-)

Table 2: Parameter estimates of AGARCH and GARCH models

Series	Method	α_0	α_1	β_1	λ	μ	γ
Apple	AGARCH	0.025	0.011	0.986	2	-0.094	0.049
		(0.014)	(0.003)	(0.004)	(-)	(0.105)	(0.019)
	GARCH	0.029	0.012	0.986	2	0.128	0
		(0.016)	(0.003)	(0.004)	(-)	(0.059)	(-)
Barclays	AGARCH	0.024	0.046	0.949	2	0.065	0.002
		(0.008)	(0.007)	(0.007)	(-)	(0.056)	(0.019)
	GARCH	0.024	0.046	0.949	2	0.071	0
		(0.008)	(0.007)	(0.007)	(-)	(0.032)	(-)
British Airways	AGARCH	0.091	0.108	0.883	2	0.038	0.015
		(0.026)	(0.015)	(0.016)	(-)	(0.064)	(0.020)
	GARCH	0.090	0.107	0.884	2	0.078	0
		(0.026)	(0.015)	(0.016)	(-)	(0.035)	(-)
Dow Jones	AGARCH	0.011	0.086	0.908	2	0.148	-0.069
		(0.003)	(0.010)	(0.011)	(-)	(0.030)	(0.020)
	GARCH	0.012	0.086	0.907	2	0.061	0
		(0.004)	(0.010)	(0.011)	(-)	(0.016)	(-)
Microsoft	AGARCH	0.036	0.079	0.919	2	-0.055	0.046
		(0.010)	(0.012)	(0.011)	(-)	(0.053)	(0.017)
	GARCH	0.035	0.075	0.923	2	0.057	0
		(0.010)	(0.011)	(0.011)	(-)	(0.033)	(-)
NASDAQ	AGARCH	0.015	0.100	0.898	2	0.210	-0.071
		(0.005)	(0.012)	(0.012)	(-)	(0.036)	(0.018)
	GARCH	0.015	0.100	0.899	2	0.092	0
		(0.005)	(0.012)	(0.012)	(-)	(0.021)	(-)
Nokia	AGARCH	0.025	0.030	0.969	2	0.374	-0.068
		(0.011)	(0.005)	(0.005)	(-)	(0.080)	(0.017)
	GARCH	0.027	0.030	0.968	2	0.134	0
		(0.011)	(0.005)	(0.005)	(-)	(0.050)	(-)
Shell	AGARCH	0.022	0.068	0.928	2	0.107	-0.010
		(0.007)	(0.011)	(0.011)	(-)	(0.045)	(0.019)
	GARCH	0.022	0.068	0.928	2	0.088	0
		(0.007)	(0.011)	(0.011)	(-)	(0.025)	(-)

probability $(1-p)$ the holder suffers a loss which is less than or equal to $-\text{VaR}$. Using the approximation $(P_{t+k}-P_t)/P_t \approx \log(P_{t+k}/P_t)$ we can write

$$p \approx \mathbb{P}[\log(P_{t+k}/P_t) \leq \text{VaR}(t,k,p)/P_t \mid \mathcal{F}_t].$$

Therefore, within an approximation, the problem of finding a VaR value is reduced to a problem of finding a conditional p^{th} quantile of the continuously compounded return $\log(P_{t+k}/P_t)$. Denote it by $q(t,k,p)$. Then $\text{VaR}(t,k,p) = e^{q(t,k,p)}P_t$. In the following we focus on estimating $q(t,k,p)$, the VaR of the log return series.

In our context loss means the decrease of price of financial asset. However, in practice the asset holder suffer loss only when she sell asset at lower price than she has bought it.

For the long position we consider the left tail of the return distribution but for the so called short position we focus on the right tail. In practice, the short position is more rarely treated than the long position.

4.2 Fitting the model

In the empirical part we study eight financial time series, daily index series NASDAQ and Dow Jones as well as daily price series of Apple, Barclays, British Airways, Microsoft, Nokia and Shell. Our samples starts at the beginning of 1995 and ends at 02/07/2006. We take into account stock splits on stock price series and construct log-returns series multiplied by 100, i.e. $y_t = 100\log(P_t/P_{t-1})$. The series have 2771–2895 observations. We have made all computations within the R environment (Ihaka and Gentleman, 1996).

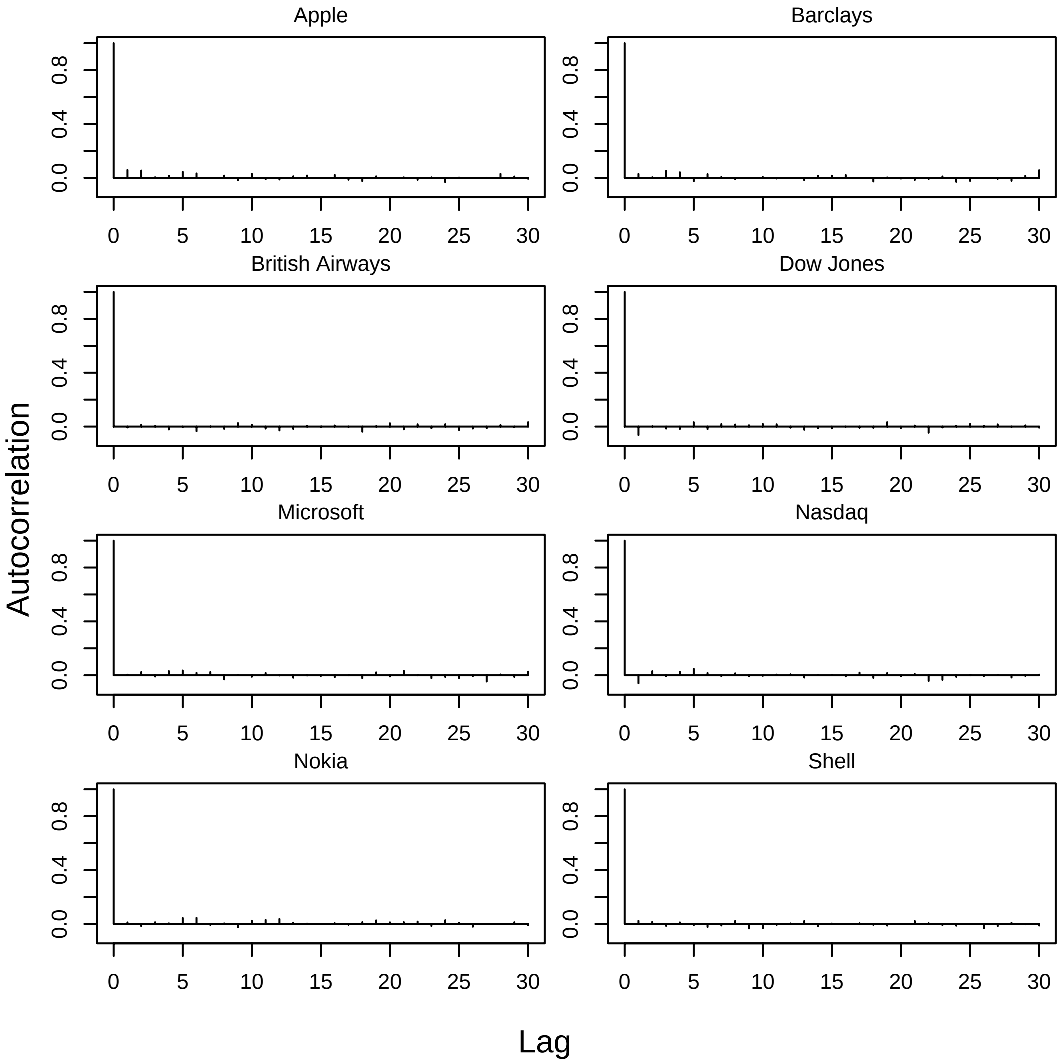

Figure 1: Autocorrelations of the ranks of squared residuals.

Table 1 gives parameter estimates and standard errors results from PEGARCH(1,1) and APEGARCH(1,1) models. In each case the estimates satisfy the restrictions $\widehat{\alpha}_0 > 0$ and $\widehat{\alpha}_1 + \widehat{\beta}_1 < 1$. We find that in stock returns series the power parameter $\widehat{\lambda}$ varies between 1.2–1.4. But in index series it is bit higher, 1.5 for Dow Jones and 1.7 for NASDAQ. There are no substantial

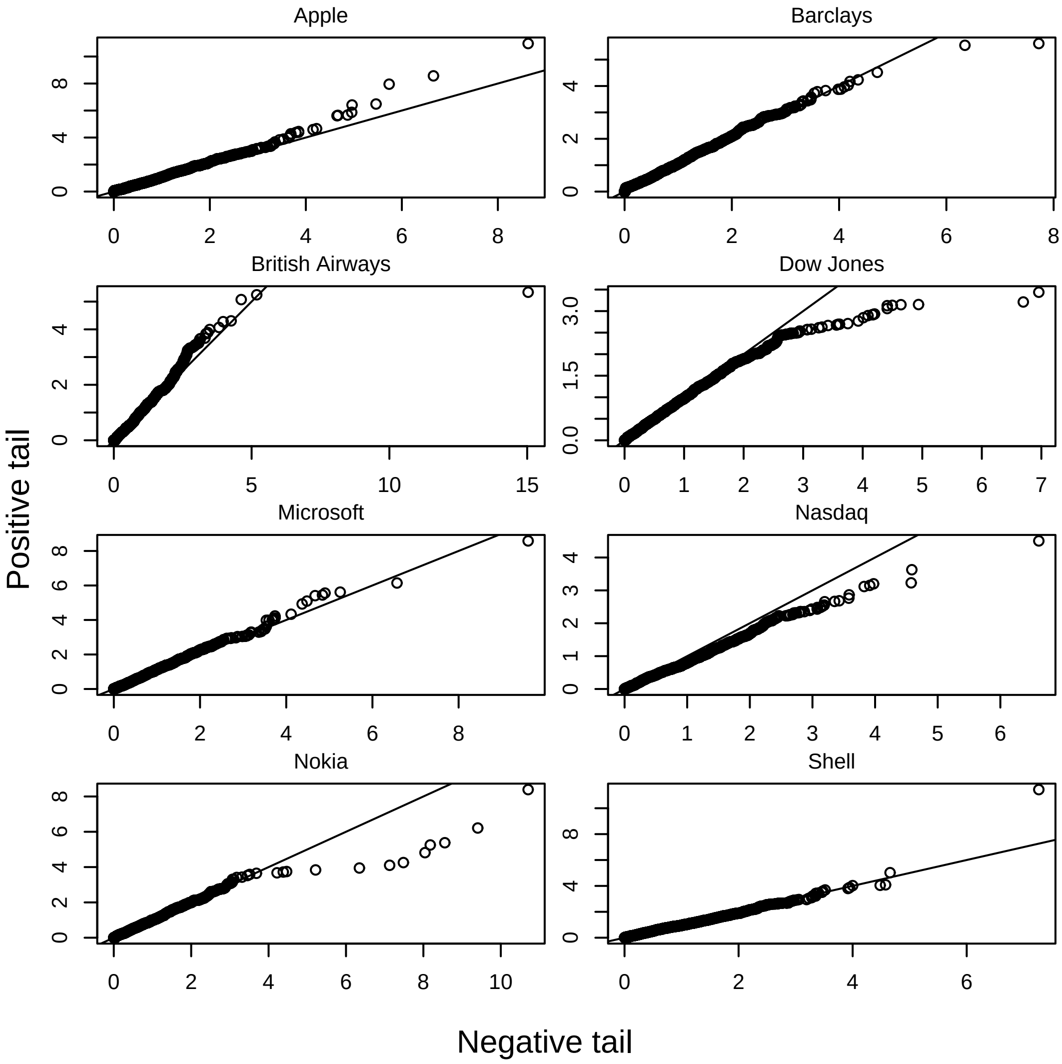

Figure 2: Symmetry plot of the residuals from the fitted PEGARCH model.

differences in $\widehat{\lambda}$ between the symmetric and asymmetric specifications of the same series. In all cases the estimates differ significantly from both 1 and 2, thus the assumption that errors ε_t are Laplace or normally distributed is rejected. A significant skewness parameter occurs in Apple and Microsoft (positive) as well as in Dow Jones and NASDAQ (negative). The corre-

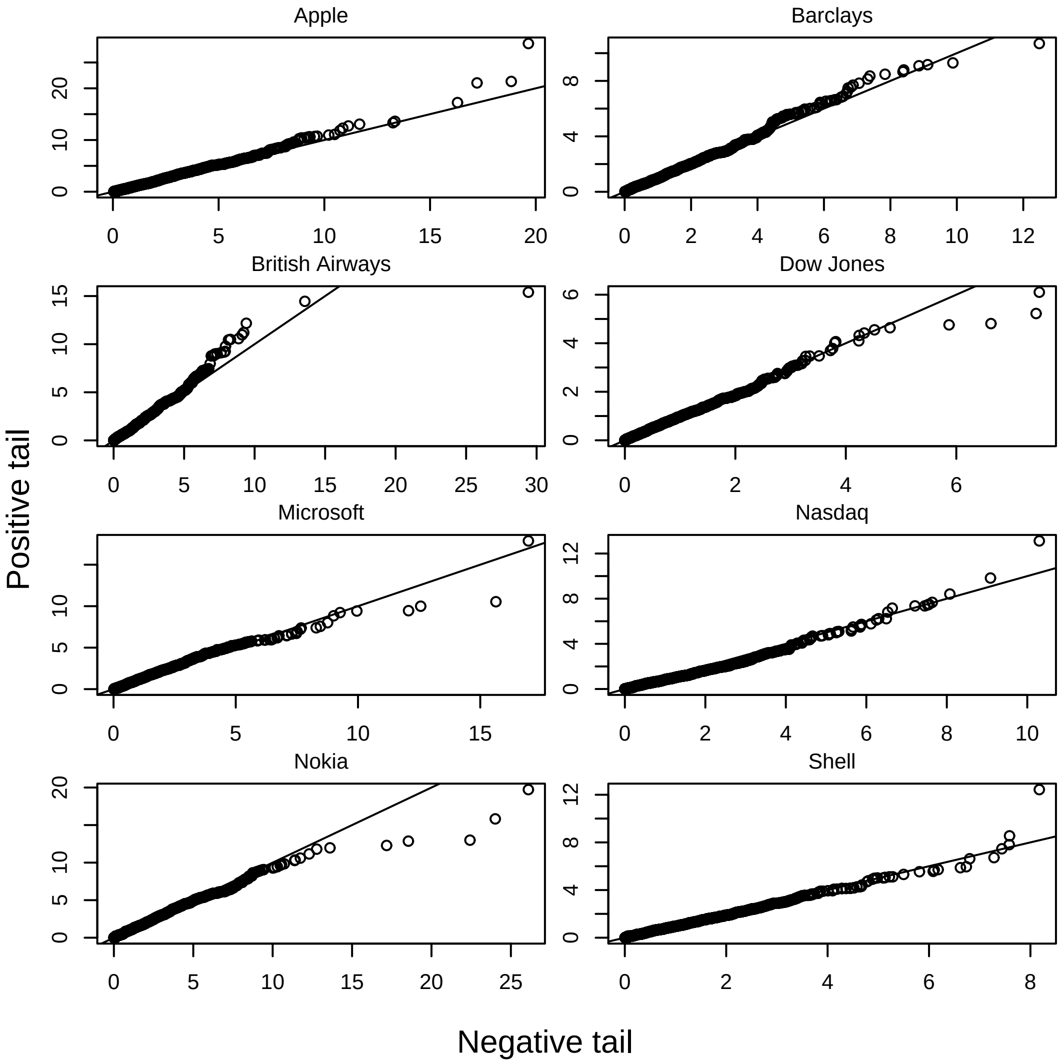

Figure 3: Symmetry plot of the return series.

sponding estimates $\widehat{\alpha}_0, \widehat{\alpha}_1, \widehat{\beta}_1$ in symmetric and asymmetric specifications of the same series are very close to each other.

In Table 2 we have the estimates from AGARCH(1,1) and GARCH(1,1) models. We find that the estimates of $\widehat{\gamma}$ have the same signs in APEGARCH and AGARCH models. Also the respective estimates $\widehat{\alpha}_0, \widehat{\alpha}_1, \widehat{\beta}_1$ are again

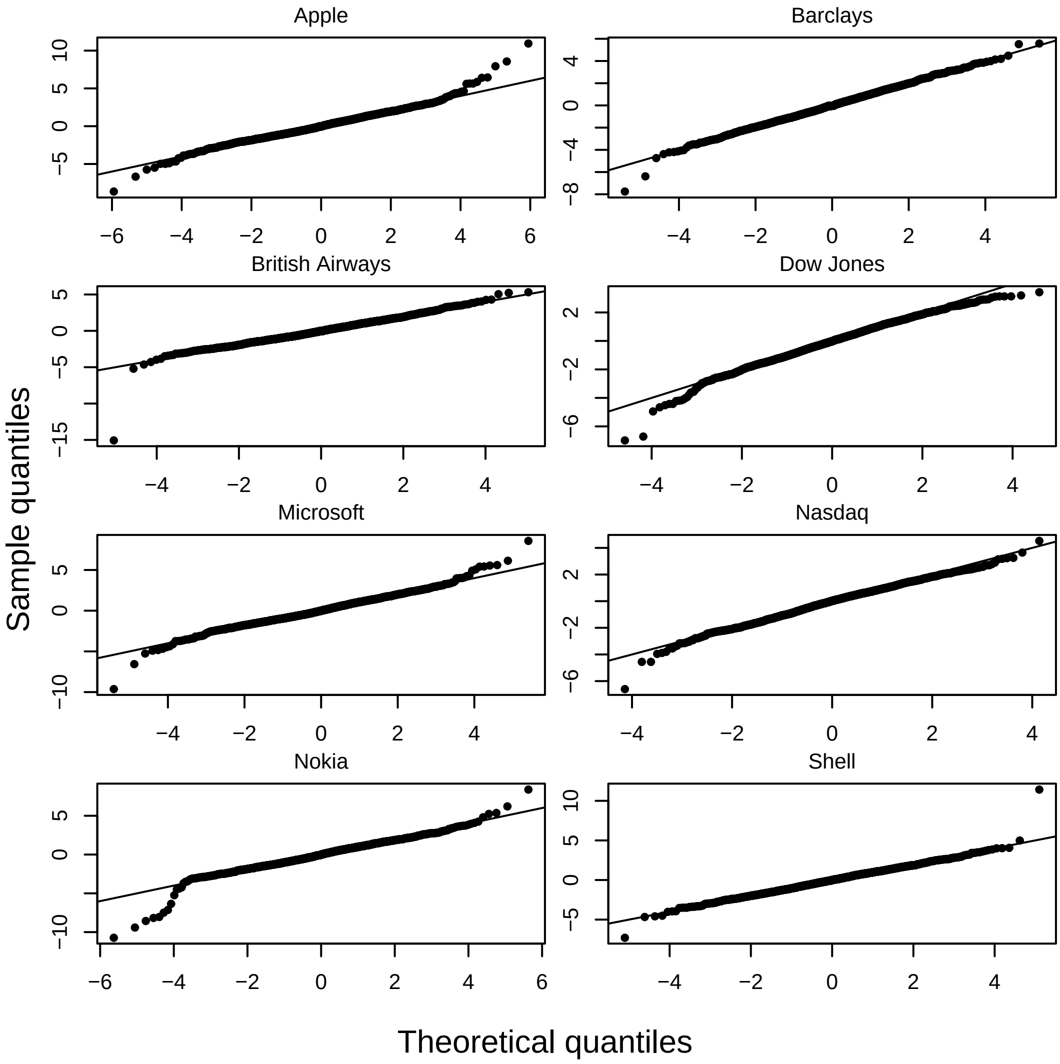

Figure 4: Quantile to quantile plot of the residuals from the fitted APE-GARCH model.

very close in AGARCH and GARCH models, though somewhat different from those in APEGARCH and PEGARCH models.

The figure 1 shows the autocorrelations function of ranks of the squared residuals. Note that the ranks remain the same whatever power of the absolute values of the residuals are used. We have used ranks because there are outliers among the residuals which may distort autocorrelations of the genuine squared residuals. As we see, only minor autocorrelation is left. The Ljung-Box test (Ljung and Box, 1978; and McLeod and Li, 1983) having an approximate χ^2 distribution is commonly used for a significance test. The appropriate number of degrees of freedom for the χ^2 distribution is not known and with the lack of a better choice we have actually used $K-p-q$ degrees of freedom for K autocorrelations when the model of order (p,q) is fitted. In actual computations we have used $K=12, p=1, q=1$. For our series the smallest p-values are for Apple (0.052) and for Nokia (0.072). The rest are well above 0.1. With the qualification concerning Apple, the fitted volatility dynamics seem to be adequate.

In figures 3–2 we exhibit plots that we call symmetry plots. Let us suppose that we have a sample from a symmetric distribution. Then with $k < n/2$, the distance from the median to the k^{th} smallest value and the distance from the k^{th} largest value to the median value should be approximately equal. We plot these values against each other, and under symmetry we expect the points lie on a straight line with an intercept zero and a slope one. Putting the negative tail is on the horizontal axis yields the interpretation that negative skewness is seen when the values are below the straight line. When the opposite is true positive skewness occurs. Figure 2 shows symmetry plots of the scaled residuals from the PEGARCH fits. We have taken the residuals from a model with assumed symmetry in order to see whether the figures suggest we should use asymmetric error distribution. In most cases we see, indeed, a clear indication of skewness. Further, the negative/positive division is the same in the plots as in in the estimates of the skewness parameters $\widehat{\gamma}$ of the APEGARCH fits. The possible exception is Shell which has $\widehat{\gamma} = -0.032$ but the corresponding plot seems fairly symmetric (e.g. compared to Nokia where $\widehat{\gamma} = -0.031$ and the plot showing a clear negative skewness). When comparing the plots of return series and the corresponding residual plot we find that in Dow Jones and NASDAQ negative skewness is more prominent in the residuals than in the return series.

In Figure 4 we have plotted the ordered scaled residuals from APEGARCH$(1,1)$ fits against the corresponding theoretical quantiles of the estimated APE distribution. The plots are called quantile to quantile plots. In all figures the black dots should lay on the straight line through origin with slope one. We find this to be true on the central part of the data, but a marked deviation is observed on the tails either generally or in terms of a few outliers. Especially, Apple, Microsoft and Nokia show more kurtosis, and Dow Jones still exhibit extra (negative) skewness. Note that the kurtosis and skewness is measured with respect to the fitted APE distribution. Apart from outliers, the residuals in others behave adequately.

4.3 Estimation of Value at Risk

Because conditionally $y_{t+1} \sim \text{APE}(\mu, \sigma_{t+1}, \lambda, \gamma)$ the VaR values for time horizon $k = 1$ with probability p are simply

$$\text{VaR} = \mu + \sigma_{t+1} z_p(\lambda, \gamma) \tag{9}$$

where $z_p(\lambda, \gamma)$ is the p^{th} quantile of $\text{APE}(0, 1, \lambda, \gamma)$. After inserting the parameter estimates the desired empirical frequency is obtained by simply checking the scaled residuals against an appropriate quantile. In our application we focus on the long position case, in fact, we have chosen $p = 0.01, 0.02, \ldots, 0.10$.

For a general time horizon k the problem is more complicated as the change in value of our financial position is $\triangle V(t,k) = y_{t+1} + \cdots + y_{t+k}$ which has an unknown conditional distribution given $\mathcal{F}_t$. A common approach, especially suitable for the GARCH model, is to find the conditional variance $\tau_t^2(k) = \mathrm{var}(\triangle V(t,k) \,|\, \mathcal{F}_t)$, and then assume that conditionally $\triangle V(t,k) \sim N(k\mu, \tau_t^2(k))$. Unfortunately, the normal assumption for $\triangle V(t,k)$ fails even for GARCH with normal errors not to speak of our more general models. Therefore, we have used simulation techniques in our calculations.

Assume for simplicity of notation that the length of the series n is even. Our experiment is performed as follows:

1. For $j = 1, \ldots, T$ with $T = \frac{n}{2} - 10$ do 2,3,4.

2. Use $y_j, \ldots, y_{j-1+\frac{n}{2}}$ to find the estimated parameters vector $\widehat{\phi}_j$.

3. Using $\widehat{\phi}_j$ compute the VaR estimates $\widehat{\mathrm{VaR}}(j-1+\frac{n}{2}, k, p)$ either from (9) (when $k = 1$) or by simulation, with 10000 replications, using recursions (3) and (4) (when $k = 5, 10$). Use values $p = 0.01, 0.02, \ldots, 0.10$.

4. For the chosen values of p and k compute the zero-one values

$$\begin{aligned} U(j,k,p) &= 1, \quad \text{if } \triangle V(j-1+\frac{n}{2}, k) < \widehat{\mathrm{VaR}}(j-1+\frac{n}{2}, k, p), \\ &= 0, \quad \text{otherwise.} \end{aligned}$$

5. Compute averages $\bar{U}(k,p) = T^{-1} \sum_j U(j,k,p)$.

The algorithm is easily changed to a bootstrap estimation method of VaR. Only the step 3 needs modification. Instead of drawing random errors from the relevant APE distribution we sample them from the scaled residuals after fitting the model to the subseries $y_j, \ldots, y_{j-1+\frac{n}{2}}$.

In Figures 5–7 we have plotted the values $\bar{U}(k,p)$ as percentages (dashed lines) versus the nominal values, for $k = 1, 5, 10$, respectively. Thus, if the model is correct we expect to see the dashed lines be close to the line with slope one. The 95 % tolerance lines are the standard limits

$$p \pm 1.96\sqrt{\frac{p(1-p)}{T}}, \quad 0.01 \leq p \leq 0.10.$$

For the case $k = 1$ the bootstrap is clearly preferable to model simulation. Remarkably, the differences between models (i.e. between estimated volatility dynamics) are negligible. Only the APEGARCH model competes well with bootstrap.

The situation with $k = 5$ in Figure 6 is somewhat worse than the one-step predictions. The asymmetric models with bootstrap seem to give similar and reliable estimates for the majority of the series. The most deviant series are Nokia and Barclays. The former is often above the upper tolerance limit whereas the latter is below the lower tolerance limit.

Apart from Nokia and Barclays the results for $k = 10$ in Figure 7 show that reasonable VaR estimates are achieved via bootstrap. Also model based AGARCH is comparable to these.

5 Discussion

In this paper we have determined and analyzed VaR measures using APE-GARCH model. This model allows for clustering volatility, asymmetric and leptokurtic behavior. By fixing parameters appropriate way the model re-

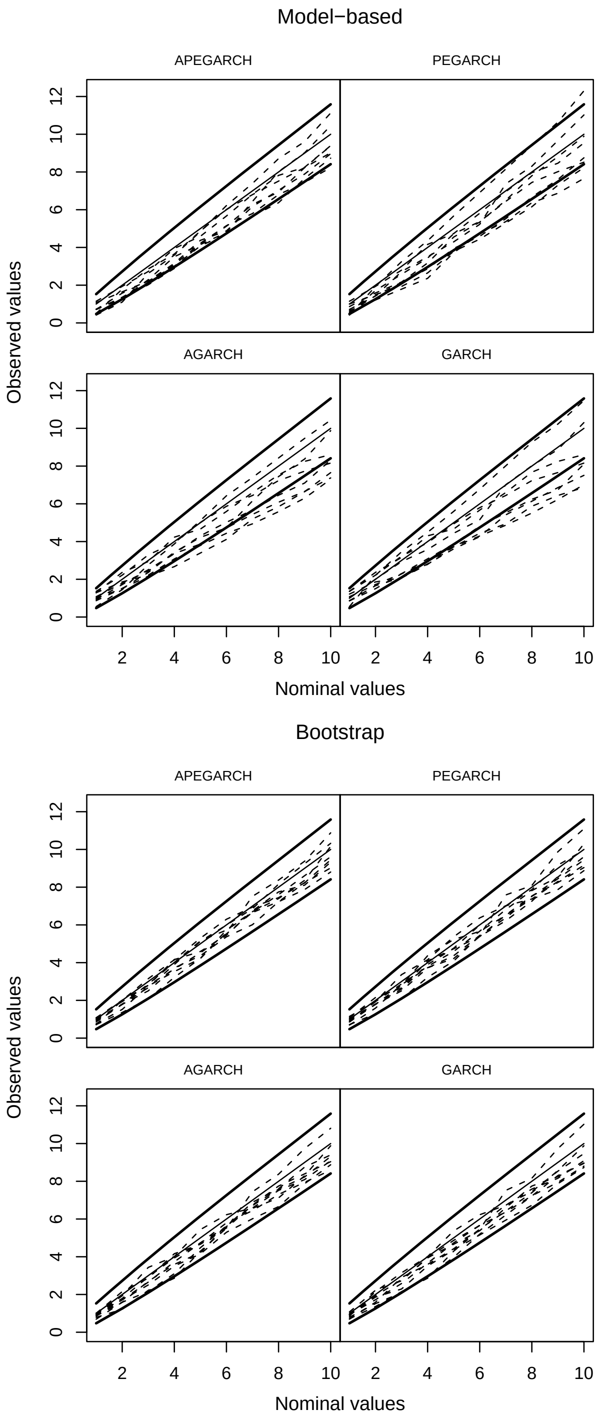

Figure 5: One-step ahead Value at Risk; dashed lines are observed values, the thick solid line in the middle is the nominal line and the other two thick lines are 95% tolerance lines

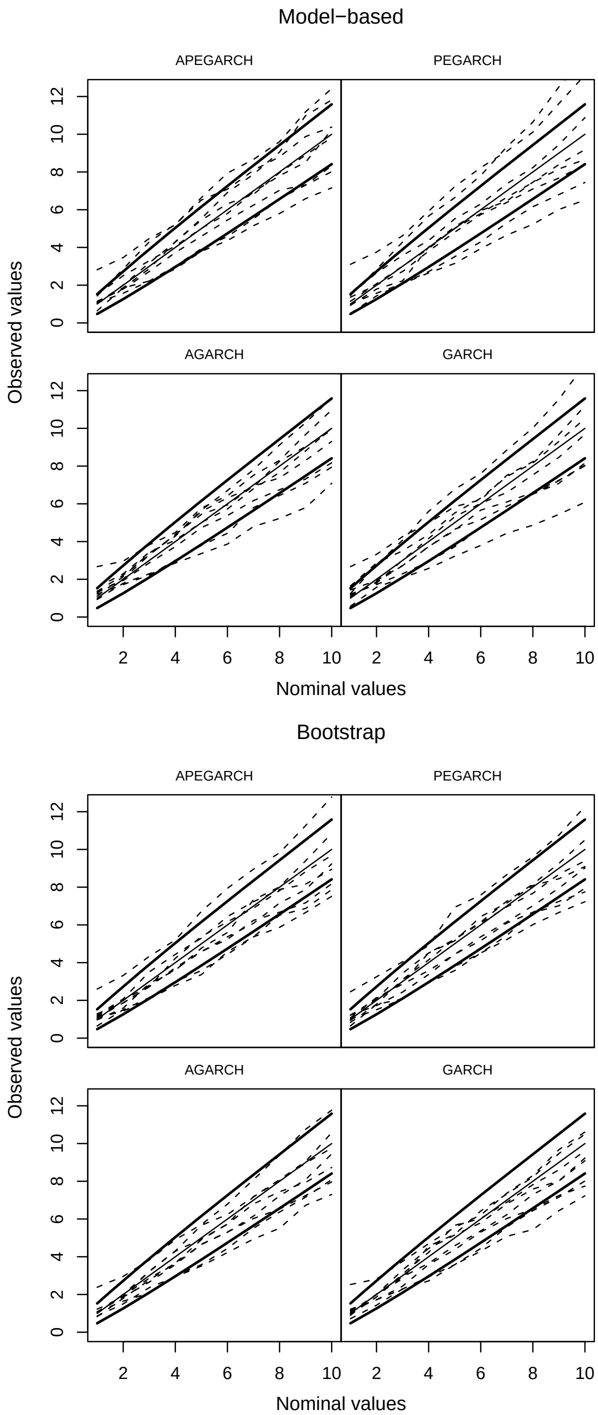

Figure 6: Five-step ahead Value at Risk; dashed lines are observed values, the thick solid line in the middle is the nominal line and the other two thick lines are 95% tolerance lines

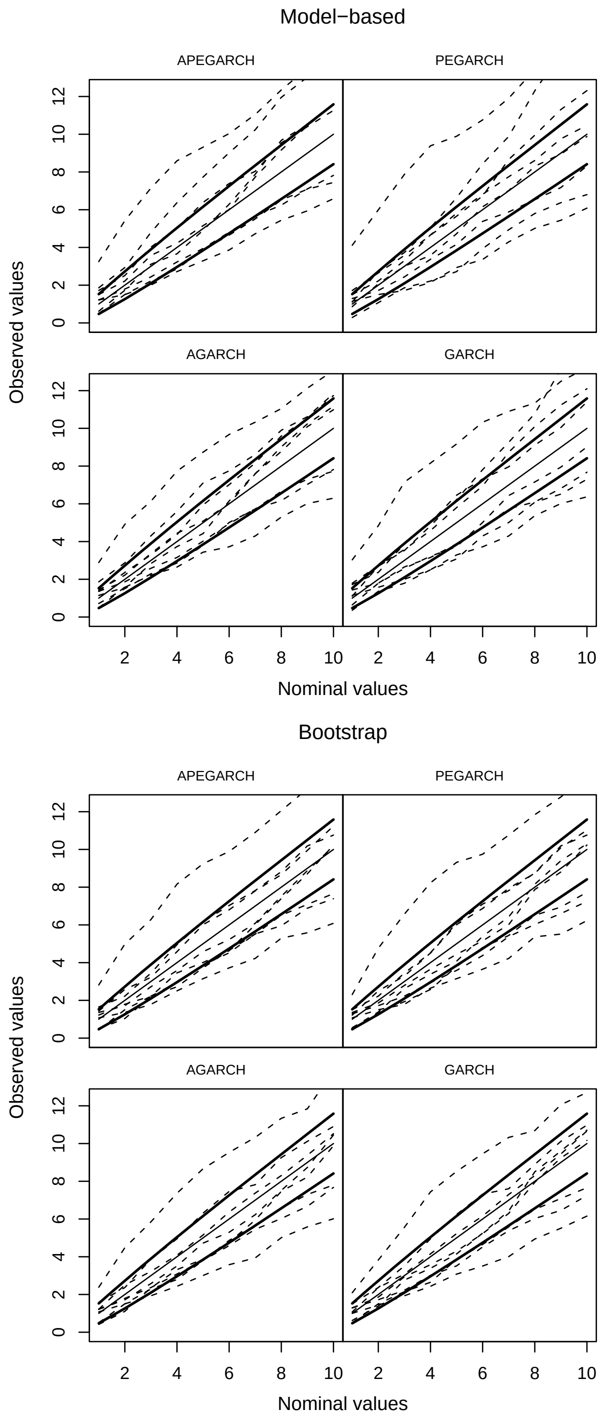

Figure 7: Ten-step ahead Value at Risk; dashed lines are observed values, the thick solid line in the middle is the nominal line and the other two thick lines are 95% tolerance lines

duces to PEGARCH, AGARCH or GARCH model. We fitted these four models to two daily stock indices and six daily stock return series and determined VaR for these series. We applied model based and bootstrap simulation techniques to calculate VaR for one, five and ten days ahead.

We found that innovations of the time series are often asymmetric and power parameter were between 1.2 and 1.7. The results indicates that APEGARCH model is a preferable model in most cases. However, the quantile to quantile plots of the residuals indicate that model is not always able to capture the tail behavior.

As a general conclusion we observe that the bootstrap simulation combined with model based volatility estimation works best in VaR calculation. The differences between models are minor though asymmetric models have some advantage. We also observe that with an increasing time horizon the VaR estimation becomes less accurate and that, when predicting 10 step ahead, estimated VaR values can be unreliable in some cases.

A Appendix

A.1 Gradients

Consider the likelihood (6) in its general form where $r_t = y_t - \mu - \gamma|y_t - \mu|$. Let θ be the vector of all other parameters than λ. Write $\log L = \sum \ell_t$. Then

$$\frac{\partial \log L}{\partial \theta} = \sum \frac{\partial \ell_t}{\partial \theta}$$

with

$$\frac{\partial \ell_t}{\partial \theta} = c_\gamma + \frac{1}{\lambda}\frac{1}{\sigma_t^\lambda}\left[\left(\frac{|r_t|^\lambda}{\sigma_t^\lambda} - 1\right)\frac{\partial \sigma_t^\lambda}{\partial \theta} - \lambda |r_t|^{\lambda-1}\operatorname{sign}(r_t)\frac{\partial r_t}{\partial \theta}\right],$$

where c_γ is a vector with zeroes apart from the coordinate corresponding to the partial derivative of $\partial \log C(\gamma, \lambda)/\partial \gamma$. Differentiation for λ yields

$$\begin{aligned}\frac{\partial \ell_t}{\partial \lambda} = \frac{1}{\lambda^2}\left[\log \lambda + \psi\left(1 + \frac{1}{\lambda}\right) - 1\right] + \frac{1}{\lambda^2}\left(\log \sigma_t^\lambda + \frac{|r_t|^\lambda}{\sigma_t^\lambda}\right) + \\ \frac{1}{\lambda}\frac{1}{\sigma_t^\lambda}\left[\left(\frac{|r_t|^\lambda}{\sigma_t^\lambda} - 1\right)\frac{\partial \sigma_t^\lambda}{\partial \lambda} - |r_t|^\lambda(\log|r_t|)\right],\end{aligned}$$

where ψ denotes the digamma function $\psi(u) = \mathrm{d}\log\Gamma(u)/\mathrm{d}u$.

The partial derivatives for σ_t^λ using (4) with $p = q = 1$ yields

$$\begin{aligned}\frac{\partial \sigma_t^\lambda}{\partial \theta} &= \frac{\partial \alpha_0}{\partial \theta} + \frac{\partial \alpha_1}{\partial \theta}|r_t|^\lambda + \alpha_1 \lambda |r_t|^{\lambda-1}\operatorname{sign}(r_t)\frac{\partial r_t}{\partial \theta} + \frac{\partial \beta_1}{\partial \theta}\sigma_{t-1}^\lambda + \beta_1\frac{\partial \sigma_{t-1}^\lambda}{\partial \theta} \\ \frac{\partial r_t}{\partial \theta} &= -\frac{\partial \mu}{\partial \theta} - \frac{\partial \gamma}{\partial \theta}|y_t - \mu| + \gamma \operatorname{sign}(y_t - \mu)\frac{\partial \mu}{\partial \theta} \\ \frac{\partial \sigma_t^\lambda}{\partial \lambda} &= \alpha_1 |r_t|^\lambda \log|r_t| + \beta_1 \frac{\partial \sigma_{t-1}^\lambda}{\partial \lambda}.\end{aligned}$$

A.2 Properties of APE distribution

It is illuminating to consider how to generate random variables from $\mathrm{APE}(0, 1, \lambda, \gamma)$. It is useful as such but as a byproduct we can establish some important properties of asymmetric power exponential distributions.

First, by the change of variable technique we find that if $Z \sim \text{PE}(0,1,\lambda)$, then $|Z|^\lambda$ follows Gamma distribution with shape $1/\lambda$ and scale λ. Thus, $E(|Z|^\lambda) = 1$. Further, let $V \sim \text{Gamma}(1/\lambda, \lambda)$ and $U \sim \text{Uniform}[0,1]$ independently from V. Then

$$Z = \frac{V^{1/\lambda}}{\text{sign}(U-(1-\gamma)/2)-\gamma} \sim APE(0,1,\lambda,\gamma).$$

Explanation is that $V^{1/\lambda}$ is distributed as the absolute value of $\text{PE}(0,1,\lambda)$ variable, and that this is multiplied by $-1/(1+\gamma)$ with probability $(1-\gamma)/2$ and by $1/(1-\gamma)$ with probability $(1+\gamma)/2$. The formula also leads straightforwardly to moments and absolute moments

$$\begin{aligned} E(Z^k) &= \frac{\Gamma\left(\frac{k+1}{\lambda}\right)\lambda^{k/\lambda}}{\Gamma\left(\frac{1}{\lambda}\right)}\left(\frac{1+\gamma}{2(1-\gamma)^k} + (-1)^k\frac{1-\gamma}{2(1+\gamma)^k}\right), \\ E(|Z|^k) &= \frac{\Gamma\left(\frac{k+1}{\lambda}\right)\lambda^{k/\lambda}}{\Gamma\left(\frac{1}{\lambda}\right)}\left(\frac{1+\gamma}{2(1-\gamma)^k} + \frac{1-\gamma}{2(1+\gamma)^k}\right). \end{aligned}$$

Therefore the mean is

$$E(Z) = m = m(\gamma,\lambda) = \frac{\Gamma(2/\lambda)\lambda^{1/\lambda}}{\Gamma(1/\lambda)}\frac{2\gamma}{1-\gamma^2}. \tag{10}$$

Note that the formula of absolute moments holds also for non-integer powers.

Write $Z-\gamma|Z| = (1-\gamma\,\text{sign}(Z))|Z|$, $Z \sim \text{APE}(0,1,\lambda,\gamma)$. Since $\text{sign}(Z) = \text{sign}(U-(1-\gamma)/2)$, we have

$$Z-\gamma|Z| = \frac{1-\gamma\,\text{sign}(U-(1-\gamma)/2)}{\text{sign}(U-(1-\gamma)/2)-\gamma}V^{1/\lambda} = \pm V^{1/\lambda}$$

which yields

$$E(|Z-\gamma|Z||^\lambda) = 1. \tag{11}$$

References

[1] Berkes, I., Horváth, L. and Kokoszka, P. (2003), "GARCH processes: Structure and estimation." *Bernoulli* 9, 201–227.

[2] Bollerslev, T. (1986), "Generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics* 31, 307–327.

[3] Bougerol, P. and Picard, N. (1992a), "Strict stationarity of generalized autoregressive processes." *The Annals of Probability* 20, 1714–1730.

[4] Bougerol, P. and Picard, N. (1992b), Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics* 52, 115–127.

[5] Ding, Z., Granger, C.W.J., and Engle, R.F. (1993), "A long memory property of stock market returns and a new model." *Journal of Empirical Finance* 1, 83–106.

[6] Giot, P. and Laurent,S. (2004), "Modelling daily Value-at-Risk using realized volatility and ARCH type models." *Journal of Empirical Finance* 11, 379–398.

[7] Ding, Z., and Granger, C.W.J. (1995), "Some Properties of Absolute Return. An Alternative Measure of Risk." *Annales d'Economie et de Statistique* 40, 67–92.

[8] He, C. and Teräsvirta, T. (1999), "Statistical properties of the asymmetric power ARCH process." In R.F. Engle and H. White (eds.) *Cointegration, causality and forecasting, Festschrift in honor of Clive W.S. Granger*. Oxford University Press: Oxford.

[9] Hull, J. and White, A. (1998), "Incorporating volatility updating into the historical simulation method for value-at-risk." *Journal of Risk* 1, 5–19.

[10] Ihaka, R. and Gentleman, R. (1996), "R: A language for data analysis and graphics." *Journal of Computational and Graphical Statistics* 5, 299–314.

[11] Jorion, P. (1997), *Value at Risk: The new benchmark for Controlling Market Risk*. McGraw-Hill: Chicago.

[12] Komunjer, I. (2007), "Asymmetric power distribution: Theory and applications to risk measurement." *Journal of Applied Econometrics* 22, 891–921.

[13] Kuester,K. Mittnik,S. Paolella,M.S. (2006), "Value-at-Risk prediction: A comparision of alternative strategies." *Journal of Financial Econometrics* 4, 53–89.

[14] Ljung, G.M. and Box, G.E.P. (1978), "On a measure of lack of fit in time series." *Biometrika* 65, 297–303.

[15] McLeod, A.I. and Li, W.K. (1983), "Diagnostic checking ARMA time series models using squared-residual autocorrelations." *Journal of Time Series Analysis* 4, 269–273.

[16] Tsay, R.S. (2005), *Analysis of Financial Time Series*, Second Edition. New York: Wiley.