NTNU
Norwegian University of
Science and Technology

Forecasting Volatility In European Equity Indices

Rune Bakken

Department of Economics

Submission date: May 2016
Supervisor: Gunnar Bårdsen, ISØ

Norwegian University of Science and Technology
Faculty of Social Sciences and Technology Management

Preface

This master's thesis is written as an ending work of my master degree program at the Norwegian University of Science and Technology (NTNU), Department of Economics.

The paper investigates the variance forecasting performance of univariate GARCH models on selected European equity indices. The topic is chosen due to personal as well as academic interest. Text and tables have been produced in Microsoft Word, while Microsoft Excel has been used for numerical and graphical analysis. Models have been estimated in G@RCH 7, an OxMetrics module.

I would like to thank my supervisor Professor Gunnar Bårdsen at the Department of Economics for his valuable advice and guidance.

The author alone is responsible for the contents and any errors.

May 2016, Trondheim

Rune Bakken

Rune Bakken

Abstract

In this thesis first order univariate GARCH models are applied to three European equity indices, DAX30, FTSE100 and OMXS30. The objective is to determine which one of the included models is best suited for out-of-sample variance forecasting while also investigating whether in-sample model fit measures provide a good indication. Models will also be fitted to symmetric and skewed Student *t* distributions, in addition to the normal distribution, to see if forecasting accuracy benefits noticeably.

The included forecasting models are: the standard GARCH model, the EGARCH model, the GJR-GARCH model, the equally weighted moving average and the exponentially weighted moving average.

Empirical results show that the normal GJR-GARCH model should be preferred for forecasting. It does not outperform in every case it but it is very reliable. Assuming a non-normal conditional distribution generally does not help on forecasting performance. Information criteria, which are used to measure in-sample fit, favor non-normal asymmetric GARCH models.

Keywords: GARCH, EGARCH, GJR-GARCH, Volatility Forecasting, DAX 30, FTSE 100, OMXS 30, Conditional Variance, EWMA

Table of Contents

List of Figures

List of Tables

1 Introduction

Financial data is often found to exhibit features that cannot be modeled using traditional econometric tools. As markets have become more electronic, information access has improved, which has made it easier to observe these features. It is now well known that the conditional variance of most financial assets follows a non-linear process that to some extent may be predicted by deterministic or stochastic models. Finance academics and practitioners alike assumed for a long time that it was random which left the main focus on returns.

Time-varying variance is especially relevant for option valuation. Realized volatility of the underlying instrument before maturity is the only unknown variable and one of the most influential ones in the widely used Black and Scholes (1973) model. With the use of options it is possible to engage in isolated bets on future volatility, but being able to forecast it, with some accuracy, also have other uses. Since the first Basel accord was enacted in 1988 banks have had to adapt to increasingly strict regulatory requirements. For the purpose of market risk management it is necessary to continuously measure potential future losses of assets or portfolios. Volatility does not equal risk directly but it is a key input in many value-at-risk models. Rational investors will also benefit as they seek to maximize expected returns relative to forecasted variance in investments. In this study the focus will be on variance forecasting for single assets, rather than portfolios, with the use of univariate GARCH models.

Robert Engle (1982) is acknowledged to be the one who started the econometric subject on autoregressive conditional heteroscedasticity (ARCH) models. It was the first parametric model created for the purpose of modeling conditional variance processes. The ARCH model has the potential to incorporate two key features, clustering and mean reversion in volatility. The generalized ARCH model, introduced by Taylor (1986) and Bollerslev (1986), has become the preferred model of these two because it is more parsimonious and avoids overfitting. Its biggest shortcoming is that it treats shocks symmetrically. Various asymmetric extensions have been proposed that remedies this. One of these that has become well known is the exponential GARCH model, proposed by

Nelson (1991). Another one is the GJR extension, named after its authors Glosten, Jagannathan and Runkle (1993).

In this study five models will be included for out-of-sample forecasting: the standard GARCH model, the EGARCH model, and the GJR-GARCH model, in addition to equally and exponentially weighted moving averages. The unconditional (and conditional) return distribution for high frequency financial data is often found to have non-zero skewness and high kurtosis, which is why GARCH models will also be fitted to non-normal conditional distributions. The data that is analyzed is comprised of daily observed indices for the equity markets in UK (FTSE100), Germany (DAX30) and Sweden (OMXS30). The sample span 20 years.

The models used to create out-of-sample forecasts are estimated from a rolling data window with constant length. After the whole sample has been used, variance forecasts are compared with realized variance estimates using loss functions. With today's regression software in-sample model fit measures are easily available after estimation. Evaluating and comparing out-of-sample forecasting performance is a stepwise procedure that is considerably more time consuming. Therefore, it is also worth taking a look at which models do well in-sample.

The rest of this thesis is structured as follows. The next section will cover prior research that is relevant for this thesis. Data sets will be presented in the third section, with descriptive statistics and preliminary specification tests. In the fourth section the model estimation procedure and forecast evaluation method will be explained. In section five included models and distributions will be presented. In section six results are reviewed for which models and distributions should be preferred based on information criteria. In the seventh section out-of-sample forecasting results are presented. Section eight concludes.

2 Literature review

GARCH models are often preferred over naïve models because of their ability to account for what is called stylized facts. In the early years they were not credited for being especially useful. In hindsight, this was because the forecast evaluation part was less well understood. Andersen and Bollerslev (1998) were among the first to address the significance of the variance estimator. Since then the importance of using appropriate evaluation measures has also received increased attention. This section will start with a description of relevant stylized facts. This is followed by a summary on forecast evaluation best practices. Lastly, results from some earlier studies are reviewed.

2.1 Stylized facts

Mandelbrot (1963) and Fama (1965) appear to have been the first to observe that some financial assets have clustered volatility and leptokurtic return distributions. These features are to some extent related. This could be explained by news often being released unevenly. Another plausible cause, suggested by Cont (2005), is investor inertia. Investors have thresholds to act and therefore need sufficient stimulus from news or past price action. Black (1976) identified another stylized fact, called the leverage effect. In equities it can be empirically observed that variance rises more following a large price fall than following a price rise of the same magnitude. It was initially believed that changes in the firm's riskiness caused this. As the price falls (rises) the debt-equity ratio increases (declines). But nowadays we know that the channel is far too small to fully explain the effect. Nevertheless, it will be referred to as the leverage effect in this study.

2.2 Forecast evaluation

The study of variance forecasting performance is made more difficult by the fact that conditional variance is not directly observable. Thus, a proxy for realized variance has to be used. The most commonly used proxy is daily squared returns. It is conditionally unbiased on the assumption that the expected return is zero, which is fair at a daily frequency. It has however been criticized for being a noisy estimator of day-by-day movements. An alternative is the high-low range proxy, initially suggested by Parkinson (1980). Relative to the daily squared return proxy it is more efficient, but the required data is not always as easy to obtain. It is not a conditionally unbiased estimator, so it needs to be scaled. The squared return proxy can be made more efficient by increasing the frequency. In the previously mentioned study by Andersen and Bollerslev (1998) standard GARCH is applied to spot exchange rates. They find that 1-day variance forecasts explain almost half the variation when measured by 5-minute returns. The drawback with using intraday data, if it can be obtained, is that the estimation procedure is more demanding.

Variance prediction errors can be evaluated in many ways. Especially loss functions have been used often in prior research. Model rankings can vary significantly between loss functions and there is a wide selection available. The suitability of several common loss functions is considered in Patton (2011). Another common method is to apply the Minzer-Zarnowitz regression, which involves regressing realized variance, measured by a proxy, on its forecast.

2.3 Prior forecasting studies

In addition to the forecast evaluation factors mentioned above, model rankings are also sensitive to forecast horizon lengths, the sample period and data frequency. Of the three indices that will be covered here FTSE100 has been researched the most on. OMXS30, on the other hand, has been given very little attention. To consider non-normal conditional distributions has become more common, judging from recent research, but overall the normal distribution is still overrepresented.

In the study by Peters (2001) various first order GARCH models are applied to FTSE100 and DAX30 daily sample data. Models are also fitted to non-normal conditional distributions, namely symmetric and skewed Student *t*. In addition to the models considered here, asymmetric power ARCH, by Ding et al. (1994), is also included. For FTSE100 the Akaike information criterion suggests that GJR-GARCH gives the best in-sample model fit. For DAX30 the APARCH model is favored, with the GJR-GARCH model in 2nd place. When it comes forecasting, results favor asymmetric models, but exactly which one should be preferred depends on the index. Assuming symmetric or skewed *t*-distributed errors result in more accurate forecasts overall.

Results from in-sample testing on FTSE100 by Alexander (2008) favor asymmetric models. The sample that is used overlaps with the one used here. Upgrading from standard GARCH to GJR- or exponential GARCH results in considerably higher log-likelihood values. Of these, the EGARCH model has the highest value. Models are also estimated with non-normal conditional distributions but the effect from this on log-likelihood values is much smaller.

In the study by Franses and Dijk (1996) first order GARCH models and a moving average are applied to weekly sample data for various European equity indices. The DAX30 index is one of them. An all-share index for the Swedish market is also included. For both indices the highest log-likelihood value is obtained with standard GARCH instead of an asymmetric specification. When it comes to forecasting, the moving average outperforms all GARCH models. Stylized facts fade on lower frequencies, so this is not surprising.

3 Statistical tests

The data that is analyzed in this study is comprised of daily observed indices for the equity markets in UK (FTSE100), Germany (DAX30) and Sweden (OMXS30). All sample data have been obtained through the Macrobond information service. The sample is exactly 20 years long, from January 1996 to December 2015. All calculations are based on natural logarithmic returns, r_t, defined by $r_t = ln\, p_t - ln\, p_{t-1}$, where p_t denotes daily price points. These are large market capitalization indices. That is, their constituents are the largest and most liquid companies in their respective markets, which make the indices investable and tradable. A liquid options market is available for each of them. Both DAX30 and OMXS30 are heavily weighted toward industrial companies. FTSE100 is much more diversified.

Table 3.1: Descriptive statistics for daily log returns

	DAX30	**FTSE100**	**OMXS30**
Observations	5069	5052	5018
Distribution [1]			
Mean	0.030538	0.010418	0.02728
Minimum	-8.8747	-9.2656	-8.5269
Maximum	10.797	9.3843	11.023
Variance	2.3351	1.4254	2.3495
Standard deviation	1.5281	1.1939	1.5328
Standard deviation (ann.) [2]	24.3275	18.9751	24.2793
Skewness	-0.12436	-0.15734	0.073324
Excess kurtosis [3]	4.0366	5.6669	3.5411
Jarque-Bera	<0.01	<0.01	<0.01

[1] All numbers, except skewness and excess kurtosis, are percentages. Distribution moment statistics are unconditional. [2] Annualized with the square root of time rule under the assumption that returns are independently and identically distributed (i.i.d.). [3] Standard measure by Pearson. Does not account for 'peakedness'. Also used in calculation of Jarque-Bera test statistic.

As each country has a different number of bank holidays throughout the year observation counts vary slightly. Formulae for statistics and tests with null and alternative hypotheses can be found in Appendix A.

It is not uncommon for financial asset return series to exhibit some lower order autocorrelation. The Ljung-Box Q-test will be applied to check for this in a quantitative way. The autocorrelation function will also be plotted to assess the presence of autocorrelation at individual lags. Obviously, conditional heteroscedasticity needs to be visible in the sample data for GARCH models to have any relevance. The ARCH LM test introduced by Engle (1982) will be used to test for this. Lastly, the sign bias test as proposed by Engle and Ng (1993) will be included to get an indication of whether positive and negative shocks of equal size have a different effect on future conditional variance.

Table 3.2: Preliminary tests on daily log returns

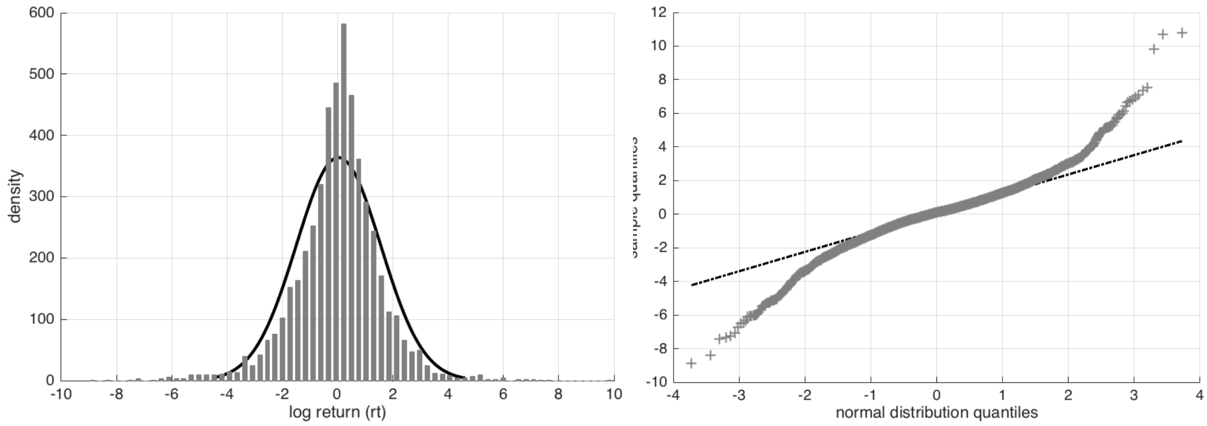
	DAX30	**FTSE100**	**OMXS30**
Ljung-Box			
Q_5	22.02	0.50	20.25
Q_{10}	50.38	1.84	28.36
Q_{20}	46.18	17.13	5.76
ARCH LM			
Q_5	<0.01	<0.01	<0.01
Q_{10}	<0.01	<0.01	<0.01
Engle-Ng			
Sign test	5.97	87.93	14.66

All numbers are probability values in percentages.

3.1 DAX30

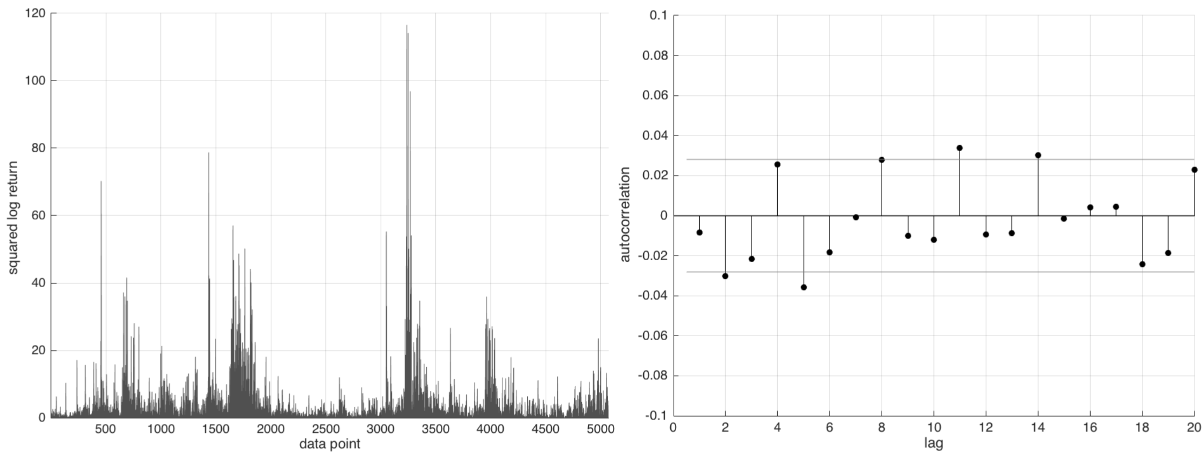
The DAX30 sample consists of 5069 daily data points. This equals approximately 253 trading days per year, on average. The annualized volatility is very high at 24.33%. The sample is not normally distributed, as can be seen in figure 3.1.1. The density plot shows how a relatively large share of observations centers on the mean. It is also heavy-tailed. The sample has more outliers than what can be expected from a normal distribution. This is illustrated in the QQ plot to the right, in which the sample distribution is fitted to a normal distribution. Many of these outliers, including the sample maximum and

Fig. 3.1.1: Sample log returns: Distribution density (left) and normal QQ plot (right)



minimum, are from late 2008. Negative skewness can also be seen in the sample distribution. Ljung-Box test results are in this case very clear. The null hypothesis of no autocorrelation can be kept for all tested lags. The joint test dilutes the significant autocorrelation coefficients that can be observed in the ACF plot at lag 5 and 11. In a general sense, it can be said that the presence of autocorrelation in DAX30 returns is insignificant. From ARCH LM test results it is evident that returns exhibit conditional heteroscedasticity. Volatility is time-varying and clustered, as can be seen in the time plot of squared returns to the left. The sign bias test indicates that positive and negative shocks have a different effect, but the significance is not especially high.

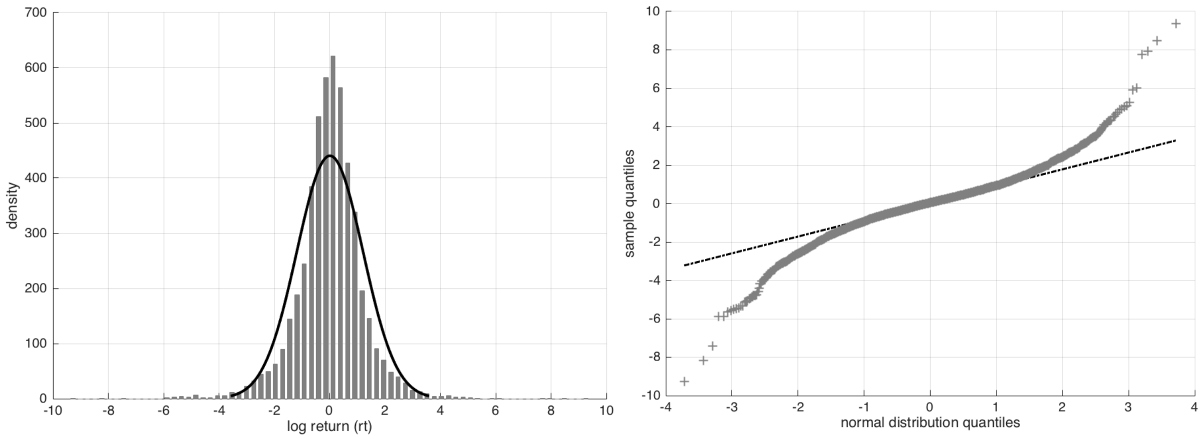
Fig. 3.1.2: Sample log returns: Squared series over time (left) and autocorrelation function (right)



3.2 FTSE100

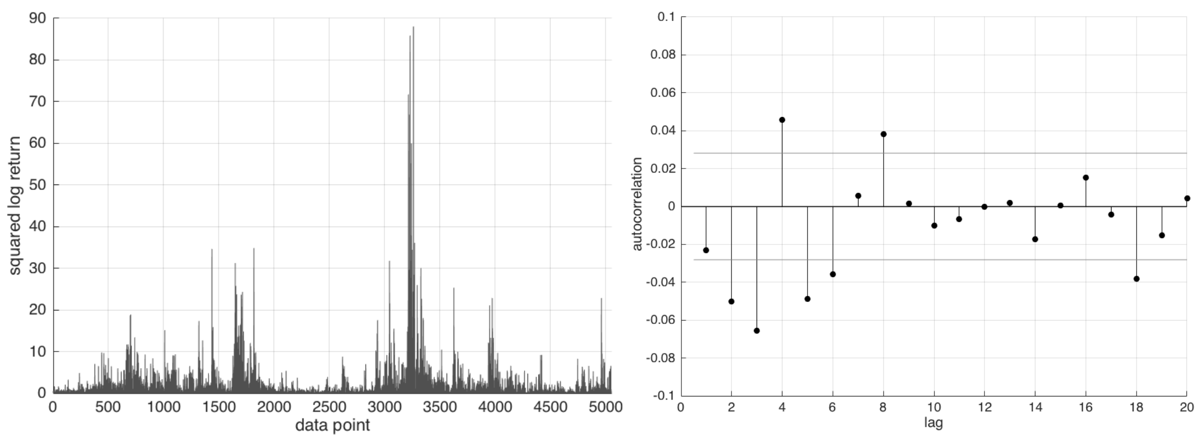
For FTSE100 there are 5052 daily data points during the sample period. This equals 253 trading days per year on average after rounding. The annualized volatility for this index is relatively low at 18.98%. Figure 3.2.1 shows that the sample is unlikely to have been drawn from a normal distribution. In the density plot to the left it can be seen that a large share of observations are of very moderate size resulting in more ‘peakedness’. Tails are also heavier than in a normal distribution. This is illustrated most effectively in the QQ plot to the right. With regards to outliers, the nine most extreme ones are all from late 2008. The sample distribution also exhibits negative skewness, but not to a great enough extent for it to be noticeable in a graphical inspection.

Fig. 3.2.1: Sample log returns: Distribution density (left) and normal QQ plot (right)



Autocorrelation is present in FTSE100 returns according to the Ljung-Box Q-test. The null hypothesis can clearly be rejected for lag 5 and 10. The sample ACF also points to this. The ACF plot shows that autocorrelation coefficients for 6 of the first 8 lags are significantly different from zero. Results of the ARCH LM test conducted on residuals controlled for autocorrelation show that the series exhibit conditional heteroscedasticity. This is easily confirmed when inspecting the time plot of squared returns to the left. Volatility is time-varying and typically occurs in bursts. No leverage effect appears to be present. The sign bias test results indicate that the effect from positive and negative shocks is similar.

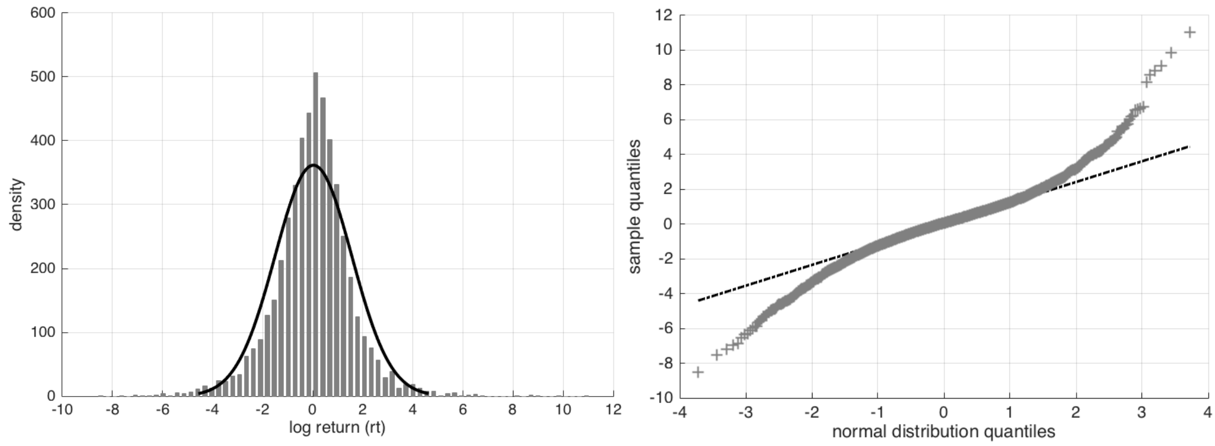
Fig. 3.2.2: Sample log returns: Squared series over time (left) and autocorrelation function (right)



3.3 OMXS30

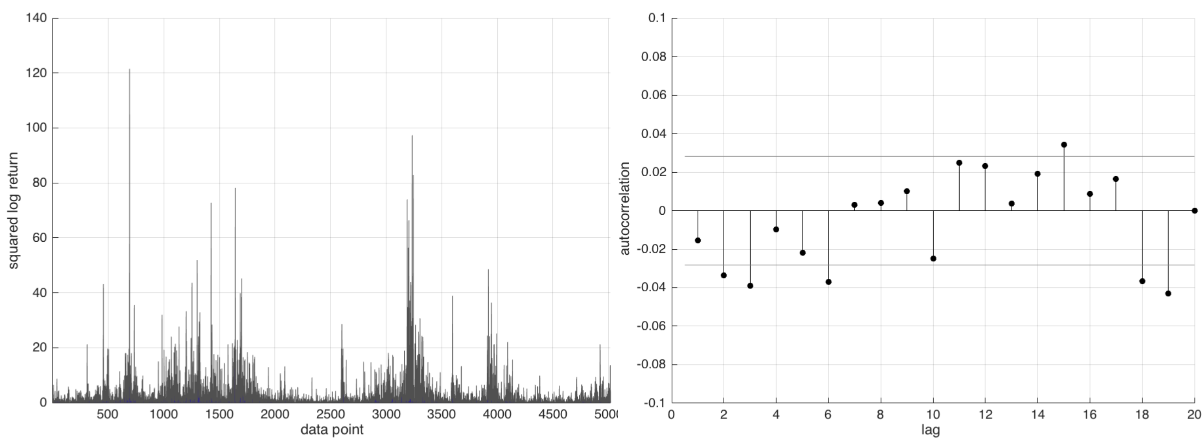
The OMXS30 sample is comprised of 5018 daily data points, which equals approximately 251 trading days per year. It has had an annualized volatility of 24.28% during the sample period. From figure 3.3.1 it can be seen that the sample is not normally distributed. A relatively large share of sample returns are modestly sized and therefore clustered around the mean, as illustrated in the density plot to the left. The sample distribution also exhibits significantly heavier tails than the normal distribution. This is illustrated most effectively by comparing the two with a QQ plot, which is shown to the right. In this case distribution skewness is positive, but too low to be deemed significant. Most of the extreme outliers are from late 2008.

Fig. 3.3.1: Sample log returns: Distribution density (left) and normal QQ plot (right)



For OMXS30 it is less obvious whether sample returns should be regarded as autocorrelated. As can be seen in the ACF plot to the right, several both lower and higher order autocorrelation coefficients are significantly different from zero.

Fig. 3.3.2: Sample log returns: Squared series over time (left) and autocorrelation function (right)



But this is not enough for the Ljung-Box Q-test to concur. The null hypothesis of no autocorrelation can be kept for all lags at a 5% significance level. The ARCH LM test confirms that returns exhibit conditional heteroscedasticity. That is, volatility is time varying and unevenly distributed. Judging from the sign bias test statistic, positive and negative shocks have a similar effect.

3.4 Summary

In this section the goal has been to become familiar with the data samples of the three indices that all further analysis will be based on. They have several characteristics in common. The return distribution of all three indices has high enough kurtosis for normality to be rejected when using the Jarque-Bera test. But this only tells half the story. They also have sharper peaks, so to speak, than a normal distribution has. FTSE100 has the most pronounced skewness, but it is not extreme in an absolute sense. It stands to reason that applying a distribution that can accommodate high excess kurtosis should result in a far better model fit for all three series, and perhaps also improved forecasting performance.

Autocorrelation statistics for the three indices vary. FTSE100 is the only index with returns that exhibit strong autocorrelation. For DAX30 and OMXS30 the ACF indicates that autocorrelation is present at some lags, but not many enough for the Ljung-Box Q-test to confirm. Specifying the conditional mean as autoregressive is likely to improve model goodness of fit, particularly for FTSE100. Forecasting performance might also benefit. There is evidence of conditional heteroscedasticity in all three indices. ARCH LM tests indicate this and it is strong enough to be clearly visible in squared return time plots. Sign bias test results are mixed. DAX30 is the only index for which positive and negative shocks of equal size appear to have a different effect on conditional variance. A leverage effect could be the cause of this.

4 Methodology

4.1 Estimation procedure

When it comes to variance forecasting the usefulness of a model is arguably best measured by how well it performs with new data. For performance rankings of different models to be meaningful forecast errors must also be measured over a long time period to obtain a large sample. To achieve this a rolling data window will be used, which can be seen as two subsamples. The first estimation window and subsample will consist of $\{r_1, r_2, \ldots, r_T\}$, where r_t denotes daily log return observations. The estimated model is used to create a conditional variance out-of-sample forecast for the period $\{r_{T+1,} r_{T+2}, \ldots, r_{T+k}\}$, the second subsample, which is then stored. The window then moves forward by k data points and the process is repeated until the whole sample has been used. Forecast errors are then computed and measured with loss functions. If N is the total number of data points there will be $\frac{\mathrm{N-T}}{\mathrm{k}}$ forecast errors (rounded down). The subsample time structure is identical for all models within each series.

Distribution attributes have a tendency to change over time so the question arises as to what is the appropriate estimation window length. It has no definite answer. There's no reason for it to be any longer than necessary. As time passes older data gradually becomes less relevant for forecasting. Moreover, withholding data increases the number of observations that can be used for out-of-sample forecasting.

In this study the estimation window length will be 10 years. Recursive estimations reveal that GARCH parameter estimates typically stabilize when the window is increased past 8 years. The equity market, and especially blue-chip companies, will be sensitive to the overall economic business cycle. A length of 10 years is likely to cover both economic expansions and contractions. The length of the forecast periods will be 10 and 20 trading days, which equals 2 and 4 weeks, respectively, if not accounting for bank holidays. The variance estimator will be based on squared daily returns

$$\hat{\sigma}_k^2 = m^{-1} \sum_{t=1}^{m} r_t^2$$

where $\hat{\sigma}_k^2$ is the realized variance estimate at time k, m is the length of the forecast period and r_t^2 represents individual log return observations. Extending the forecast past just a few days will even out idiosyncratic movements in the proxy, resulting in less ambiguous loss function results.

4.2 Forecast evaluation

Estimated and forecasted variance will be compared with the use of loss functions. The choice of loss functions can have a significant impact on model rankings. Even with all the previous research no consensus has formed yet on this subject. It should be partly determined based on which variance proxy is used. In this study the chosen methodology for forecast evaluation has been influenced by Patton (2011), and Patton and Sheppard (2009). They point out that the chosen evaluation methods should be robust to the presence of noise in the variance proxy, meaning that it should not affect how the models are ranked as the forecast error sample increases in size. Secondly, evaluation techniques should not rely on assumptions about the third and fourth moment of the conditional return distribution.

Mean squared error (MSE) and Quasi-likelihood (QLIKE) are loss functions that fit these criteria. The latter one is the preferred choice in the studies mentioned above. A lower loss function value indicates higher forecasting accuracy. They are given equal weight when models are ranked in section 7 but this might not always be appropriate. If forecasting accuracy during periods of high volatility is more important to the practitioner the MSE function should be given more weight.

4.2.1 Mean Squared Error

The optimal MSE forecast is the conditional variance so the function is appropriate. It has however been disregarded in many previous studies because extreme outliers can drive the results. But trying to counteract this by opting for functions that uses absolute or median values will instead lead to biased results. MSE is defined by

$$\mathrm{MSE} = \mathrm{n}^{-1} \sum_{\mathrm{i}=1}^{\mathrm{n}} (\hat{\sigma}_{\mathrm{i}}{}^{2} - \mathrm{h}_{\mathrm{i}})^{2}$$

Where $\hat{\sigma}_i^2$ and h_i is the conditional variance estimator and forecast, respectively, and n is the size of the forecast error sample.

4.2.2 Quasi-Likelihood

This is a standardized measure, which makes the outliers much less prominent. In similar fashion, the optimal QLIKE forecast is the conditional variance. The function is given by

$$\mathrm{QLIKE} = n^{-1} \sum_{i=1}^{n} \log h_i + \frac{\hat{\sigma}_i{}^2}{h_i}$$

It is the loss function implied in the Gaussian likelihood function. Bollerslev et al. (1994) appears to have been one of the first to consider it. Since then it has been infrequently used.

5 Forecasting models

Many new univariate GARCH models have been suggested since the standard one by Bollerslev (1986). Most of these are only slightly different from each other and will therefore produce similar results. In this case the data exhibits volatility clustering and leptokurtosis. The data can also be expected to show leverage effects, which will be explored further in Section 6. Some asymmetric models will therefore also be included. Compared with naïve alternatives GARCH models are more cumbersome to estimate, as they require regression software. The relative usefulness of naïve models is for this reason also of interest.

5.1 The Maximum Likelihood Method

Once the GARCH models have been specified the parameters will be estimated using the Quasi-Maximum Likelihood Estimator (QMLE) in G@RCH 7, an OxMetrics module. The method works by finding the most likely function that explains the observed data based on a given error distribution. It's easier to use the natural logarithm of the function, called the log-likelihood function (LLF). The specification depends on which distribution is used and its density function. The default algorithm in G@RCH 7 is the quasi-Newton method of Broyden, Fletcher, Goldfarb and Shanno (BFGS). Formulas for the respective log-likelihood functions can be found in Appendix B.

5.2 Conditional mean

The conditional mean is of secondary interest but taking it into account is nevertheless needed for correct specification. The standard model will have a constant mean μ and can be written as

$$r_t = \mu + \varepsilon_t$$

r_t represents the daily log return and ε_t denotes the market shock, error or innovation. In cases where returns are autocorrelated making the conditional mean autoregressive should improve the model fit. An AR(1) mean equation can be defined by

$$r_t = \mu + \phi r_{t-1} + \varepsilon_t$$

An AR(1) specification is likely to be sufficient to model the dynamic structure where it exists. Higher orders will therefore not be considered in this study.

5.3 Densities

How the GARCH error process, represented by ε_t, is defined has ramifications for the estimation procedure. It has zero expected value and time varying conditional variance. To assume that if follows a conditionally normal distribution is the most practical. Though, for high frequency financial data this assumption is often less realistic. This also seems to be the case here. To accommodate non-zero skewness and high kurtosis different conditional distributions will be considered, including the Gaussian or normal distribution and the symmetric and skewed Student *t* distributions.

5.3.1 The normal distribution

By assuming that errors are normally distributed, $\varepsilon_t \sim N(0, \sigma_t^2)$, skewness and excess kurtosis remain fixed at zero. The normal distribution is used extensively because it has nice properties and is easy to apply. Its standardized density function is defined as

$$\varphi(\varepsilon) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(\varepsilon-\mu)^2}{2\sigma^2}}, \quad -\infty < \varepsilon < \infty$$

5.3.2 The symmetric Student *t* distribution

When errors are assumed to follow a symmetric Student t distribution, $\varepsilon_t \sim t(0, \sigma_t^2, v)$, the degree of freedom, represented by v, is treated as a parameter and optimized to a constant level along with other parameter values. If $v = \infty$, it collapses to the normal distribution. As the degree of freedom is reduced tails become heavier while the rest of the density is lowered. The standardized Student *t* density function is defined as

$$\varphi(\varepsilon|\mathrm{v}) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{\pi(v-2)}} \left[1 + \frac{\varepsilon^2}{v-2}\right]^{-\frac{(v+1)}{2}}, \quad -\infty < \varepsilon < \infty, \; v < 2 \leq \infty$$

where $\Gamma(\cdot)$ is the gamma function.

5.3.3 The skewed Student *t* distribution

By assuming that errors follow a skewed Student t distribution, $\varepsilon_t \sim t_{sk}(0, \sigma_t^2, \xi, v)$, the asymmetry parameter is also optimized. This distribution nests the two above. It becomes symmetric if the asymmetry, represented by ξ, is set to zero. If, in addition,

$v = \infty$, it becomes normal. The standardized skewed Student t density function is given by

$$\varphi(\varepsilon|v,\xi) = \mathrm{b}\frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\pi(v-2)}\,\Gamma\left(\frac{v}{2}\right)}\left(1+\frac{\zeta^2}{v-2}\right)^{-\frac{(v+1)}{2}}, \qquad \begin{array}{c} -\infty < \varepsilon < \infty \\ v < 2 \leq \infty \\ -1 < \xi < 1 \end{array}$$

where

$$\zeta = \begin{cases} (\mathrm{bz}+\mathrm{a})/(1-\xi) \text{ if } \mathrm{z} < -\mathrm{a}/\mathrm{b}, \\ (\mathrm{bz}+\mathrm{a})/(1-\xi) \text{ if } \mathrm{z} \geq -\mathrm{a}/\mathrm{b}. \end{cases}$$

The constant terms a and b are defined as

$$\mathrm{a} = 4\xi\mathrm{c}\frac{v-2}{v-1}, \qquad \mathrm{b}^2 = 1 + 3\xi^2 - \mathrm{a}^2$$

and are introduced to obtain a variable ε with zero mean and unit variance, while c given by

$$\mathrm{c} = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\pi(v-2)}\,\Gamma\left(\frac{v}{2}\right)}$$

5.4 Conditional variance

The conditional variance is the variable that will be focused on in this study. Based on statistical test results in section 3 the GARCH framework should provide superior predictive capability compared to naïve models. The various models that are included in this study will be presented in the following subsections with conditional variance specifications. Model definitions and forecasting formulas are similar to those given by Alexander (2008). The number of parameter lags used will be limited to one, ie. $p,q = 1$. In most cases that is sufficient to capture the conditional dependencies. GJR-GARCH and EGARCH will be included to investigate the possibility of asymmetric news responses. Obviously, there are more complex specifications available that might be better able to capture the various asymmetric effects. The problem is that chances of convergence problems increase as more parameters are introduced. Since hundreds of models will be estimated for variance forecasting, this is an important issue.

5.4.1 Equally weighted moving average

The specification is identical to that of the conditional variance proxy that uses squared returns on a daily timeframe. Even though this is a relatively simple forecast model it is still widely used today. The procedure works by using the current reading as a constant forecast for the whole forecast period. In this case, when applied to forecasting, it will be called the simple moving average (SMA) to make it more distinctive. The average will have twice the length to that of the forecast period. That is, the 10-day forecast will be based on the previous 20 observations while the 20-day forecast will be based on 40 observations. The SMA's greatest weakness is perhaps its inability to adjust for mean reverting variance. Unless the forecast period is very short, it will cause forecast errors to often be large when the conditional variance is deviating far from its long-run average. Equal weighting also slows down the model's response time. Furthermore, when previous outliers drop out abrupt fluctuations will occur, irrespective of current developments.

5.4.2 Exponentially weighted moving average

When trying to predict future volatility fresh data is the most useful. In the EWMA model older observations gradually lose importance. This study will follow the weighting scheme outlined in Alexander (2008)

$$h_t = (1-\lambda)\varepsilon_{t-1}^2 + \lambda h_{t-1}$$

where h_t denotes the conditional variance. The value of λ, called lambda, decides how quickly the forecast adjusts to new innovations. The optimal value will vary depending on the data set and the length of the forecast period. The 1-step ahead forecast based on all information available up to and including observation T is

$$h_{T+1} = (1-\lambda)\varepsilon_T^2 + \lambda h_T$$

The EWMA model is effectively a restricted version of the symmetric GARCH model, but it has the same problem as the SMA model. The constant ω is set to zero and the other parameter values sum to 1, so the forecast will be the same for all time horizons.

5.4.3 The GARCH model

The symmetric GARCH model as proposed by Bollerslev (1986) is considerably more powerful than the naïve models introduced earlier because of its ability to capture conditional heteroscedasticity and mean reversion. With one lag the conditional variance equation can be written as

$$h_t = \omega + \alpha\,\varepsilon_{t-1}^2 + \beta\,h_{t-1}$$

Parameter constraints might be needed to ensure that the conditional and unconditional variance is positive and finite. They can be written together as

$$\omega > 0, \qquad \alpha, \beta \geq 0, \qquad \alpha + \beta < 0$$

The GARCH model has created much more practical interest than the ARCH model proposed by Engle (1982) because it is more parsimonious and easier to estimate. It does not treat positive and negative shocks differently, which should put it at a disadvantage when applied to equities.

The sum of α and β tells how quickly the conditional and unconditional variances converge. Together with the constant ω the unconditional variance, or long-run average, can be obtained

$$\bar{h} = \frac{\omega}{1 - (\alpha + \beta)}$$

Since shocks have a symmetric effect the conditional variance given by this model features a 'floor'. This also leads to a higher unconditional variance than in asymmetric models.

Conditional variance forecasts are derived recursively. Together they form a variance term structure that slopes upward (downward) if the current conditional variance is below (above) the long-run average. The 1-step ahead forecast based on all information available up to and including observation T is

$$h_{T+1} = \omega + \alpha\, \varepsilon_T^2 + \beta\, h_T$$

All the RHS values are in this case known. Taking the expectation of the unknown squared error ε_{T+1}^2 the forecast for step 2 at time T can now be obtained

$$h_{T+2} = \omega + \alpha\, E_k(\varepsilon_{T+1}^2) + \beta\, h_{T+1} = \omega + (\alpha + \beta) h_{T+1}$$

This procedure is repeated for the rest of the term structure. In general, for S>1, the 1-step ahead forecast at $T+S$ is

$$h_{T+S+1} = \omega + (\alpha + \beta) h_{T+S}$$

When the forecast period to be evaluated consists of more than one day the average of the term structure during that period is used.

5.4.4 The GJR-GARCH model

The impact from shocks on conditional variance in the standard GARCH model is symmetric. For some financial assets this specification is inadequate. Glosten et al. (1993) proposed a modification to capture a potential asymmetric effect. With one lag this model takes the form

$$h_t = \omega + \alpha\, \varepsilon_{t-1}^2 + \beta\, h_{t-1} + \delta\, I_{\{\varepsilon_{t-1}<0\}} \varepsilon_{t-1}^2$$

where the indicator function $I_{\{\varepsilon_{t-1}<0\}} = 1$ if $\varepsilon_t < 0$ and 0 otherwise. To ensure that the conditional variance is positive the following parameter constraints are imposed

$$\omega > 0, \quad \alpha, \beta, \delta \geq 0$$

The coefficient δ will be positive if negative innovations have a greater effect on conditional variance. Assuming that the distribution of ε_t is symmetric around zero, the expectation of the asymmetry parameter is

$$E\left[I_{\{\varepsilon_{T-1}<0\}} \varepsilon_{T-1}^2\right] = \frac{1}{2} \delta h_T$$

The long-run average is then given by

$$\bar{h} = \frac{\omega}{1 - (\alpha + \beta + \frac{1}{2}\delta)}$$

The forecasting procedure is similar to that of the standard GARCH model. The starting forecast based on all information available up to and including observation T is

$$h_{T+1} = \omega + \alpha\, \varepsilon_T^2 + \beta\, h_T + \delta\, I_{\{\varepsilon_T<0\}} \varepsilon_T^2$$

This can then be used at time T to compute a forecast for step 2

$$h_{T+2} = \omega + \alpha\, E_T[\varepsilon_{T+1}^2] + \beta\, h_{T+1} + \delta\, E_T\left[I_{\{\varepsilon_T+1<0\}} \varepsilon_{T+1}^2\right] = \omega + \left(\alpha + \beta + \frac{1}{2}\delta\right) h_{T+1}$$

The rest of the term structure can be derived in a similar way. At observation *T+S,* for *S*>1, the 1-step ahead forecast is

$$h_{T+S+1} = \omega + (\alpha + \beta + \frac{1}{2}\delta)\, h_{T+S}$$

5.4.5 The EGARCH model

The exponential GARCH model will be the most complex one of those included in this study. In addition to the standard features it allows for asymmetric effects between positive and negative errors. With one parameter lag the model can be written as

$$\ln(h_t) = \omega + \beta \ln(h_{t-1}) + g(z_{t-1})$$

where $g(z_{t-1})$ is the asymmetric response function

$$g(z_t) = \alpha z_t + \gamma[|z_t| - E|z_t|]$$

Negative shocks have an impact $\alpha - \gamma$ on the log of the unconditional variance while for positive shocks the impact is $\alpha + \gamma$. z_t represents the normalized random variable

$$z_t = \frac{\varepsilon_t}{\sqrt{h_t}}$$

The value of $E|z_t|$ depends on the assumption made on the unconditional density of z_t. If assuming that it belongs to a normal distribution, the value is simply $\sqrt{2/\pi}$. If it follows a skewed Student t distribution the value is given by

$$E|z_t| = \frac{4\xi^2}{\xi + 1/\xi} \frac{\Gamma\left(\frac{1+v}{2}\right)\sqrt{v-2}}{\sqrt{\pi}\Gamma(v/2)}$$

where ξ is the asymmetry parameter and v denotes the degree of freedom of the distribution. If z_t instead follows a symmetric Student t distribution, then $\xi = 1$. By instead taking the logarithm of the unconditional variance parameter constraints become unnecessary. h_t will stay positive even with negative parameter values.

The unconditional variance for exponential GARCH is defined by

$$\bar{h} = \exp\left(\frac{\omega}{1-\beta}\right)$$

The 1-step ahead forecast at time T is relatively straightforward, as all the right-side quantities are known

$$h_{T+1} = \exp(\omega)\exp(g(z_T))h_T^{\beta}$$

Further out the forecast will rely on the conditional expectation at origin T. The forecast for step 2 is given by

$$h_{T+2} = \exp(\omega)\, E_T[\exp(g(z_{T+1}))]h_{T+1}^{\beta}$$

The prior expectation of the response function is obtained by

$$E[\exp(g(z)] = \int_{-\infty}^{\infty} \exp[\alpha z + \gamma(|z| - E|z|)]\,\varphi(z)dz$$

$$= \exp(-\gamma E|z|)\left[\int_0^{\infty} \exp\big((\alpha+\gamma)z_t\big)\,\varphi(z)dz + \int_{-\infty}^{0} \exp((\alpha-\gamma)z)\varphi(z)dz\right]$$

$$= \exp(-\gamma E|z|)\left[\exp\left(\frac{1}{2}(\gamma+\alpha)^2\right)\phi(\gamma+\alpha) + \exp\left(\frac{1}{2}(\gamma-\alpha)^2\right)\phi(\gamma-\alpha)\right]$$

φ and ϕ denote the PDF and CDF of the assumed error distribution. Consequently, the full formula for the 2-step ahead forecast is

$$\mathrm{h}_{\mathrm{T+2}} = \left[\exp\left(\frac{1}{2}(\gamma+\alpha)^2\right)\phi\,(\gamma+\alpha) + \exp\left(\frac{1}{2}(\gamma-\alpha)^2\right)\phi(\gamma-\alpha)\right]\exp(\omega - \gamma\,\mathrm{E}|\mathrm{z}|)\mathrm{h}_{\mathrm{T+1}}^{\beta}$$

The rest of the term structure can be derived in a similar way. The 1-step ahead forecast at observation *T+S* is given by

$$\mathrm{h}_{\mathrm{T+S+1}} = \exp(\omega - \gamma\,\mathrm{E}|\mathrm{z}|)\,\mathrm{h}_{\mathrm{T+S}}^{\beta}\left[\exp\left(\frac{1}{2}(\gamma+\alpha)^2\right)\phi\,(\gamma+\alpha) + \exp\left(\frac{1}{2}(\gamma-\alpha)^2\right)\phi(\gamma-\alpha)\right]$$

6 In-sample analysis

Before moving on the empirical results of variance forecasting the sample data will be examined by taking a look at which types of models perform well in-sample. Previous studies on in-sample modeling have been reviewed earlier. Results in this case can be expected to be similar. Three models will be presented for each index, one for each GARCH type. These are chosen based on Schwarz and Hannan-Quinn information criteria. A lower criterion indicates a better goodness of fit.

$$\mathrm{SIC} = -2\frac{\mathrm{Log\ L}}{\mathrm{n}} + 2\frac{\log(\mathrm{k})}{\mathrm{n}}, \quad \mathrm{HQIC} = -2\frac{\mathrm{Log\ L}}{\mathrm{n}} + 2\frac{\mathrm{k}\log[\log(\mathrm{n})]}{\mathrm{n}}$$

Log L denotes the maximized LLF value, k is the number of parameters and n is the number of observations used in the estimation.

The adequacy of the fit will be evaluated with statistical diagnostics on the model's standardized residuals, $\varepsilon_t/\sqrt{h_t}$. A correctly specified model will produce standardized residuals that fit the specified error distribution. In addition, they should not display any autocorrelation or conditional heteroscedasticity. The procedure will be similar to that in Section 3. The Ljung-Box portmanteau test will be used to check for autocorrelation while the ARCH LM test will show if conditional heteroscedasticity is still present. Details on how these tests are conducted can be found in Appendix A. Determining whether the distribution is appropriate is more complicated. The main objective in this paper is not to ascertain which models perform best based on in-sample evaluation. Some of the estimated models will have error distributions that are optimized with regards to both the third and fourth moment. The complexity could be increased past this to obtain an even better fit, for example with time varying distribution parameters, but this will be left to other researchers.

Models were estimated using 10 years of data. Specifically, from 1st January 2001 to 31st December 2010. Using the full sample would result in very rigid models with flat variance term structures. The estimation period is the center of the sample, which for all three indices include two bear markets. Between these two periods markets were unusually calm. The numbers in parentheses are t-statistics. Coefficients are rounded to four decimals. Extreme outliers will not be managed in any way to improve model fit.

6.1 DAX30

The models with the lowest information criteria have an asymmetric specification. Switching from standard GARCH to either GJR or EGARCH results in a noteworthy improvement in model fit. The exponential GARCH model has here a slightly higher log-likelihood value than the GJR-GARCH model but when the number of parameters is taken into account they score about equally well. None of the presented models has an autoregressive conditional mean. Including a first order AR parameter results in only marginal increases in log-likelihood values while absolute t-ratios indicate low significance.

The highest ranking models have conditional error distributions with high kurtosis and negative skewness, but the fitted values are not unusual. Unconditional volatilities are for the most part realistic. As is often the case, standard GARCH has the highest long-run volatility. For the exponential GARCH model it is too low.

Tests conducted on standardized residuals reveal that not all of the models are correctly specified. Specifically, ARCH effects are still visible in the standardized residuals belonging to the GJR-GARCH model. This could be remedied by increasing the model order. This would also reduce information criteria values further. In the same model the symmetric shock parameter value is constrained to zero, which makes the leverage parameter the only transmitter of past shocks. The sample period used in these estimations extends across several different market environments. Therefore it is inevitable that the standardized residuals for these models contain some outliers, even if they are relatively complex.

Table 6.1.1: DAX30 models, January 2001 to December 2010

	GARCH(1,1) Skewed Student *t*	GJR-GARCH(1,1) Skewed Student *t*	EGARCH(1,1) Skewed Student *t*
Log-likelihood	-4308.16	-4261.33	-4258.99
Uncond. volatility [1]	1.7025	1.3820	0.8196
Coefficients (t-stat)			
μ	0.0671 (3.01)	0.0224 (1.01)	0.0223 (1.04)
ω	0.0161 (2.80)	0.0238 (4.12)	-0.7457 (-2.08)
α	0.0895 (7.04)	0.0000 (0.00)	0.0745 (0.29)
β	0.9050 (76.46)	0.9088 (74.87)	0.9813 (253.60)
δ		0.1574 (6.93)	
γ_1			-0.1221 (-3.60)
γ_2			0.1166 (5.06)
ε distr. [2]	v 12.2895	v 20.4835	v 18.8111
	ξ -0.1014	ξ -0.1316	ξ -0.1284
Misspecification tests [3]			
Ljung-Box			
Q_5	23.22	20.34	12.95
Q_{10}	30.06	23.74	16.23
Q_{20}	40.91	28.02	19.41
ARCH LM			
Q_5	3.75	0.28	7.79
Q_{10}	27.04	3.75	18.04
Information criteria			
SIC	3.4054	3.3717	3.3729
HQIC	3.3966	3.3614	3.3612

[1] Square root of variance. [2] v and ξ represent degrees of freedom and asymmetry, respectively. [3] All numbers are p-values, Q_m denotes the number of lags

6.2 FTSE100

The two asymmetric models can be seen to provide markedly lower information criteria than the standard GARCH model. Skewed *t* EGARCH is the highest ranking model. It has the highest log-likelihood value and both information criteria are lower than for the other two models. In this case, it appears that modeling with an autoregressive mean process is appropriate. Log-likelihood values and t-ratios indicate that a first order AR parameter has significant explanatory power. This is in line with what autocorrelation statistics for the full sample indicated in section 3.

For all three specifications the lowest information criteria are obtained by assuming a conditional error distribution with high kurtosis and negative skewness. The unconditional volatility levels of the presented models are realistic with one exception. For the exponential GARCH model it is much too low.

Misspecification test results generally support the case that these models have an adequate fit. Judging from Ljung-Box test results, one autoregressive lag is sufficient to model the autocorrelation that is present in returns. However, some residual ARCH effects are present in the standardized residuals belonging to the GJR and standard GARCH model. Though the extent of this is not severe. Increasing the model order would take care of this and cause a further reduction in information criteria values. But it is also likely that the model would become overfitted.

Table 6.2.1: FTSE100 models, January 2001 to December 2010

	AR(1)-GARCH(1,1) Skewed Student *t*	AR(1)-GJR-GARCH(1,1) Skewed Student *t*	AR(1)-EGARCH(1,1) Skewed Student *t*
Log-likelihood	-3667.28	-3619.45	-3611.64
Uncond. volatility [1]	1.3623	1.0741	0.5662
Coefficients (t-stat)			
μ	0.0399 (2.61)	0.0085 (0.55)	0.0049 (0.32)
Φ	-0.0822 (-4.05)	-0.0738 (-3.79)	-0.0622 (-3.19)
ω	0.0120 (3.16)	0.0157 (4.04)	-1.7207 (-3.50)
α	0.1057 (7.68)	0.0000 (0.00)	0.0292 (0.13)
β	0.8878 (64.84)	0.9052 (67.96)	0.9849 (298.10)
δ		-0.1611 (-5.48)	
γ_1			-0.1270 (-4.61)
γ_2			0.1590 (5.39)
ε distr. [2]	v 15.7323	v 21.2052	v 23.0931
	ξ -0.1391	ξ -0.1611	ξ -0.1590
Misspecification tests [3]			
Ljung-Box			
Q_5	31.51	45.21	45.52
Q_{10}	75.60	83.96	79.23
Q_{20}	72.02	82.48	77.22
ARCH LM			
Q_5	12.48	10.73	32.40
Q_{10}	6.78	5.45	11.13
Q_{20}	1.69	5.10	9.21
Information criteria			
SIC	2.9196	2.8849	2.8818
HQIC	2.9093	2.8731	2.8686

[1] Square root of variance. [2] v and ξ represent degrees of freedom and asymmetry, respectively. [3] All numbers are p-values, Q_m denotes the number of lags

6.3 OMXS30

The two asymmetric models have the lowest information criteria. But if comparing with the other two indices, the difference is smaller. Between these two asymmetric models the ranking is mixed. The exponential GARCH model has a higher log-likelihood value but also a greater number of parameters which it is penalized for by information criteria. In this case goodness of fit is not improved noticeably when modeling with an autoregressive mean process. With regards to standard GARCH models, absolute t-ratios are almost high enough to indicate relevance at a 5% significance level.

Assuming a non-normal conditional error distribution results in lower information criteria for all three model types. The optimal *t* distributions that are assumed here have very high kurtosis. When it comes to skewness the increases in log-likelihood values are so marginal that it mostly results in higher information criteria. Specifically, for the two asymmetric models the SIC and HQIC disagrees on whether the distribution should be symmetric or not. In such cases the simpler model should be the preferred one.
Exponential and standard GARCH models typically have the lowest and highest conditional volatilities, respectively. This is also the case here, which is reflected by unconditional volatility levels. All three levels appear reasonable.

Misspecification test results suggest that these models have an adequate fit. There is little sign of residual autocorrelation or conditional heteroscedasticity in the standardized residuals of any of the models. The asymmetric models appear to have some redundant parameters. Most noteworthy is the symmetric shock parameter in the GJR-GARCH model. With alpha constrained to zero the leverage parameter is the only channel for past shocks.

Table 6.3.1: OMXS30 models, January 2001 to December 2010

	GARCH(1,1) Symmetric Student *t*	GJR-GARCH(1,1) Symmetric Student *t*	EGARCH(1,1) Skewed Student *t*
Log-likelihood	-4368.51	-4334.01	-4332.82
Uncond. volatility [1]	1.8882	1.4373	1.1494
Coefficients (t-stat)			
μ	0.0792 (3.53)	0.0390 (1.71)	0.0334 (1.46)
ω	0.0130 (2.10)	0.0183 (2.36)	0.4110 (1.82)
α	0.0747 (4.97)	0.0000 (0.00)	-0.1579 (-0.83)
β	0.9217 (60.45)	0.9277 (44.70)	0.9852 (238.40)
δ		0.1268 (4.73)	
γ_1			-0.1423 (-4.47)
γ_2			0.1297 (4.94)
ε distr. [2]	v 9.7985	v 12.6951	v 12.4722
Misspecification tests [3]			
Ljung-Box			
Q_5	50.49	28.77	18.83
Q_{10}	52.61	35.86	22.82
Q_{20}	88.84	77.46	62.67
ARCH LM			
Q_5	10.39	69.19	43.34
Q_{10}	34.95	33.46	37.63
Q_{20}	16.72	44.70	30.89
Information criteria			
SIC	3.4971	3.4707	3.4741
HQIC	3.4882	3.4618	3.4622

[1] Square root of variance. [2] v and ξ represent degrees of freedom and asymmetry, respectively. [3] All numbers are p-values, Q_m denotes the number of lags

6.4 Summary

In this section nine different GARCH variations have been estimated using 10 years of sample data. The models were ranked based on information criteria and the highest ranking one for each specification was chosen for further analysis. This step is not essential but it does allow one to learn more about the data before moving on to forecast estimation and evaluation.

The results were generally in line with previous studies. Information criteria measure goodness of fit. In this regard, asymmetric response functions made a big difference. Asymmetric models delivered lower information criteria for all three indices. When distribution kurtosis and skewness become additional parameters that can be optimized the model fit is improved further. In other words, the preferred models are asymmetric and fitted to a Student *t* distribution, symmetrical or skewed. This is in line with observations given in section 3 about the unconditional distribution for the full sample period.

It appears that with the given estimation period first order models are sufficiently complex to capture the conditional heteroscedasticity. In some cases lower information criteria could be obtained by applying second order models. However, increasing model complexity does not typically benefit its robustness. Given a relatively volatile estimation period the result quickly becomes an overfitted model if the focus is solely on obtaining standardized residuals that are well behaved.

The unconditional, or steady state, variance level differs depending on the GARCH model. It is determined indirectly from optimized parameters. Using the previously listed definitions, the variation between models in unconditional volatility for some indices is noteworthy. The unconditional volatility levels for these models are also rather high, in an absolute sense. This has to do with a volatile estimation period.

These models and observations might seem to have limited usefulness outside this section since they are not based on the full sample. But the general observations are still highly relevant. Irrespective of which estimation period is used, asymmetric models with a symmetric or skewed *t* distribution have the lowest information criteria. Log-likelihood values, coefficients and standard errors are the only details that change.

7 Forecasting results

In this section forecasting results of all the considered models will be presented. Some complications were encountered, which will be discussed first. It would seem appropriate to accommodate potential time varying autocorrelation. The preliminary Ljung-Box Q-tests conducted on returns were inadequate for discovering this. All GARCH models have therefore been estimated both with and without an AR(1) conditional mean for all three indices. The two categories should not be seen as competitors. With regards to conditional variance specifications, all models will be of first order. That is, they will have one ARCH and GARCH parameter.

When estimating GJR-GARCH models it became necessary to impose a positivity constraint on the symmetric ARCH parameter to ensure a positive conditional variance. No constraints were necessary for standard GARCH models. Lack of convergence was a frequent problem when estimating exponential GARCH models. This is rarely an issue under normal circumstances. Switching to the Sequential Quadratic Programming (SQP) algorithm helped in some cases. Working around it by excluding the occurrence from the estimation procedure is tedious and time consuming. GARCH models for which the estimation procedure had several no convergence occurrences are listed with N/A values.

Daily variances are converted to standard deviations, which are referred to as volatilities. Taking the square root will increase the usefulness of charts by shrinking the scales. The focus will mostly be on the best performing models. These will also be used to gain more insight about the population by examining changes in parameter values.
The assumed conditional kurtosis in Student *t* models, represented by the degree of freedom v, is also liable to change during the estimation procedure. As mentioned in section 5.3, the value of v is optimized along with other parameters values. Degree of freedom parameter values will be charted to see how pronounced the conditional kurtosis is and to make model rankings more insightful.

The lowest value in each column is marked with a solid underline. The second lowest value, belonging to a model with a different variance equation than the first, is marked with a dotted underline.

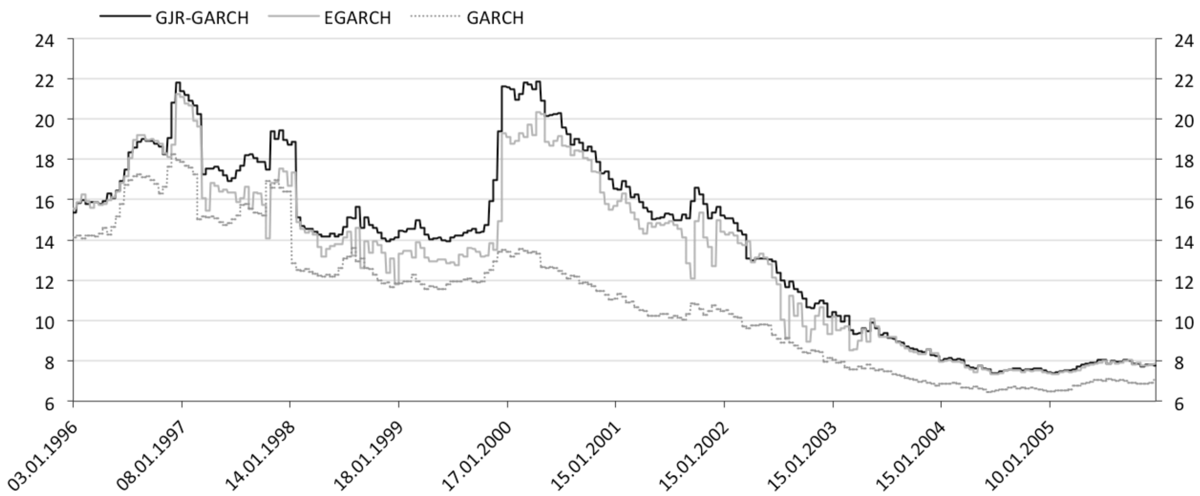
7.1 DAX30

Table 7.1.1: DAX30 out-of-sample forecasting results

	10 days		20 days	
	MSE	QLIKE	MSE	QLIKE
SMA	4.9182	1.5674	6.0898	1.6084
EWMA 0.92	4,0576	1,5324	3,7778	1,5738
EWMA 0.95	4,6379	1,5262	4,3867	1,5594
Const. in cond. mean				
GARCH normal	3.7699	1.4831	3.3938	1.5019
GARCH symmetric t	3.8987	1.4860	3.5792	1.5074
GARCH skewed t	3.8807	1.4876	3.5459	1.5081
GJR-GARCH normal	3.2858	1.4681	2.7626	1.5008
GJR-GARCH symmetric t	3.3209	1.4704	2.7721	1.5091
GJR-GARCH skewed t	3.4017	1.4690	2.9249	1.5027
EGARCH normal	3.7366	1.4790	3.5446	1.5276
EGARCH symmetric t	3.6733	1.4793	3.4372	1.5287
EGARCH skewed t	3.9084	1.4966	4.0915	1.5885
Const. + lag in cond. mean				
GARCH normal	3.7461	1.4827	3.3648	1.5015
GARCH symmetric t	3.8836	1.4856	3.5604	1.5073
GARCH skewed t	3.8542	1.4870	3.5132	1.5080
GJR-GARCH normal	3.2973	1.4682	2.7738	1.5010
GJR-GARCH symmetric t	3.3347	1.4705	2.7885	1.5059
GJR-GARCH skewed t	3.4344	1.4691	2.9674	1.5032
EGARCH normal	3.8084	1.4749	3.5955	1.5171
EGARCH symmetric t	3.6575	1.4807	3.4299	1.5293
EGARCH skewed t	N/A	N/A	N/A	N/A

Models with a symmetrical or skewed Student t error distribution have similar performance overall. Staying with the normal distribution results in the lowest loss function values for both GJR and standard GARCH models. This contradicts the observations in section 3 about the unconditional return distribution. Student t conditional distributions are also heavy-tailed, as can be seen in figure 7.1.1. The degree of freedom varies considerably during the first years. In the second half it starts on a gradual decline as the volatility settles down. The trend eventually carries it below 8 for all three models, resulting in very heavy-tailed conditional distributions.

Fig. 7.1.1: Symmetric t GARCH(1,1) models, degree of freedom parameter value over time
The horizontal axis shows the estimation window starting point



Asymmetric models stand far apart when comparing loss function values. Exponential GARCH models have some of the highest values, in contrast to GJR-GARCH models, which have some of the lowest ones. In fact, skewed t EGARCH has the worst performance, second only to naïve models. Standard GARCH models perform better with an autoregressive conditional mean. GJR-GARCH models do not. With exponential GARCH models the results are mixed in this regard. Overall, the normal GJR-GARCH model is the best performing model. It has the lowest loss function values with a significant margin, for both MSE and QLIKE, regardless of forecast length. Figure 7.1.2 shows how forecasted volatility from this model compares with estimated volatility on a 10-day horizon.

The conditional variance equation in the GJR-GARCH model has only four parameters. Figure 7.1.3 illustrates the evolution in parameter values during the sample period. The variance constant and beta is relatively stable throughout the whole sample. For alpha and delta the variation is higher. Something interesting occurs when the rolling window

Fig. 7.1.2: Normal GJR-GARCH(1,1), 10-day horizon realized (proxy) and forecasted volatility

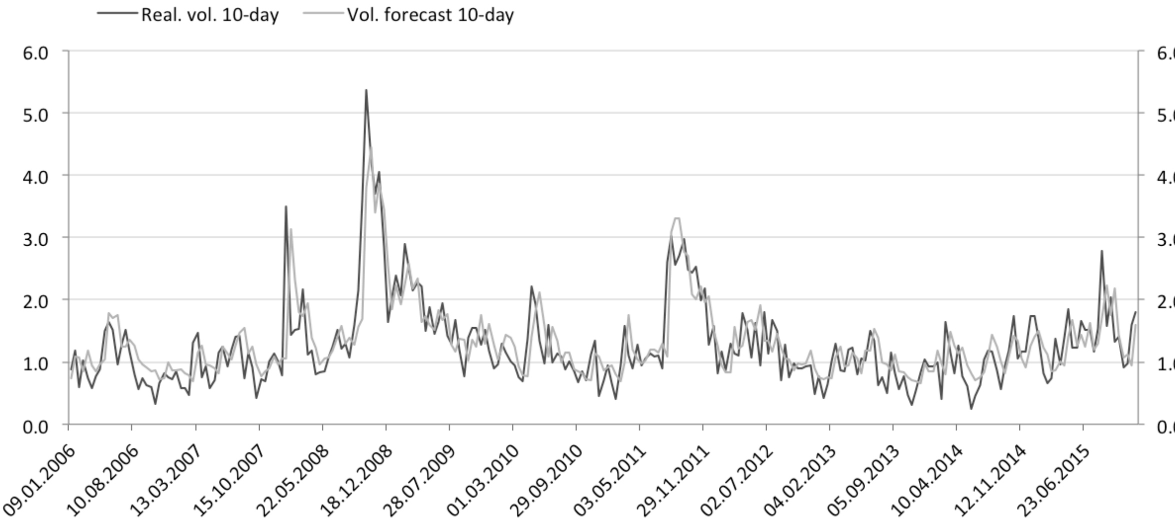
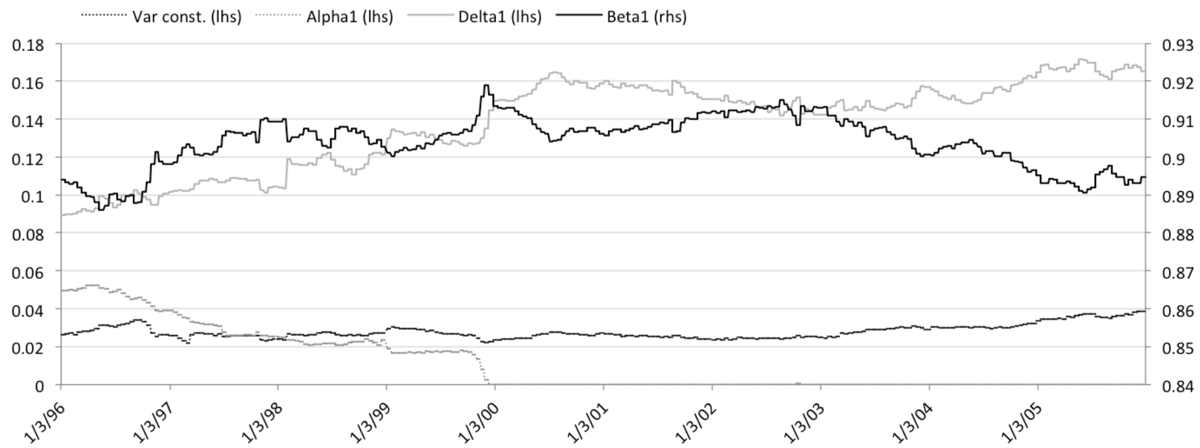


Fig. 7.1.3: Normal GJR-GARCH(1,1), model coefficients over time
The horizontal axis shows the estimation window starting date



starting point approaches year 2000. Delta jumps while alpha reaches zero after a steady downtrend. The leverage parameter essentially takes over and becomes the only parameter that responds to market shocks. If excluding models with a GJR specification, normal or symmetric *t* GARCH is the second best model for a 10-day forecast horizon. If extending it to 20 days, a standard normal GARCH model has it beat.

Moving on to naïve models. These have the lowest accuracy of all estimated models. The SMA has provides the worst forecasts. The difference relative to EWMA models is noteworthy.

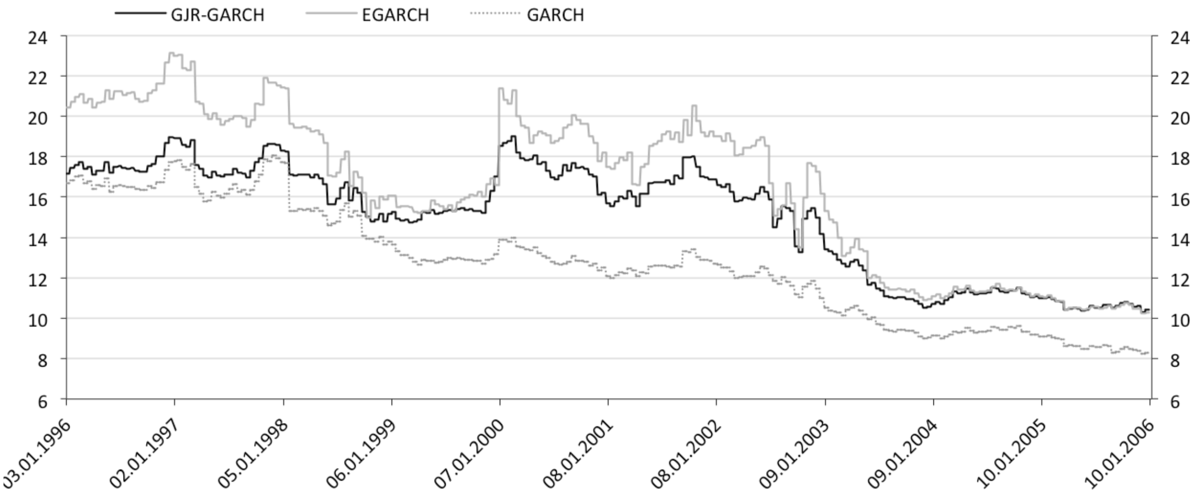
7.2 FTSE100

Table 7.2.1: FTSE100 out-of-sample forecasting results

	10 days		20 days	
	MSE	QLIKE	MSE	QLIKE
SMA	6.5897	1.2246	6.4794	1.2140
EWMA 0.92	6,2577	1,1574	4,7105	1,2000
EWMA 0.95	6.1960	1.1520	5.0350	1.1809
Const. in cond. mean				
GARCH normal	6.5135	1.1204	4.4747	1.1363
GARCH symmetric t	6.4453	1.1194	4.4878	1.1361
GARCH skewed t	6.3869	1.1219	4.4456	1.1377
GJR-GARCH normal	6.0306	1.0881	4.0732	1.1316
GJR-GARCH symmetric t	6.0636	1.0884	4.0883	1.1349
GJR-GARCH skewed t	6.2298	1.0863	4.1750	1.1318
EGARCH normal	5.5145	1.0959	4.3877	1.1490
EGARCH symmetric t	5.5191	1.0991	4.4118	1.1579
EGARCH skewed t	5.6411	1.1177	4.8718	1.2158
Const. + lag in cond. mean				
GARCH normal	6.3517	1.1169	4.4176	1.1332
GARCH symmetric t	6.4149	1.1177	4.5109	1.1346
GARCH skewed t	6.3320	1.1202	4.4672	1.1362
GJR-GARCH normal	5.9778	1.0882	4.0270	1.1318
GJR-GARCH symmetric t	6.0015	1.0886	4.0325	1.1351
GJR-GARCH skewed t	6.1516	1.0867	4.1544	1.1318
EGARCH normal	5.5355	1.0969	4.3970	1.1503
EGARCH symmetric t	N/A	N/A	N/A	N/A
EGARCH skewed t	N/A	N/A	N/A	N/A

The FTSE100 unconditional return distribution exhibits very high kurtosis. This was observed in section 3. Student *t* conditional distributions share this feature, reflected by low degree of freedom parameter values, which is illustrated in fig. 7.2.1. Nonetheless, choosing a normal distribution results in lower loss function values for GJR and exponential GARCH models. Results for standard GARCH are more ambiguous in this regard. Symmetric and skewed *t* models perform about equally well overall.

Fig. 7.2.1: Symmetric *t* GARCH models, degree of freedom parameter value over time
The horizontal axis shows the estimation window starting point



Standard GARCH is in most cases beat by asymmetric models. Although with a 20-day horizon the difference is less pronounced. GJR-GARCH models have fairly low loss function values overall. Larger variations can be seen with respect to exponential GARCH models. They have some of the highest and lowest loss function values. Skewed *t* GARCH have high enough values to be in the same league as EWMA models.

With a 10-day forecast horizon the results are mixed. The QLIKE function prefers the skewed *t* GJR-GARCH, while the MSE function favors the normal EGARCH model. The latter one manages to predict some of the outliers well, which puts it far ahead of other models when squared errors are used for evaluation. Figure 7.2.2 shows 10-day volatility forecasts generated by this model compared with estimated volatility.

If extending the forecast period to 20 days the normal GJR-GARCH model has the best performance according to both loss functions, although with different conditional mean specifications. Volatility forecasts generated by this model can be seen in figure 7.2.3. Thus, with the lowest values in three of four columns GJR models should be preferred for forecasting FTSE100 variance, chiefly with a normal error distribution.

Fig. 7.2.2: Normal EGARCH(1,1), 10-day horizon realized (proxy) and forecasted volatility

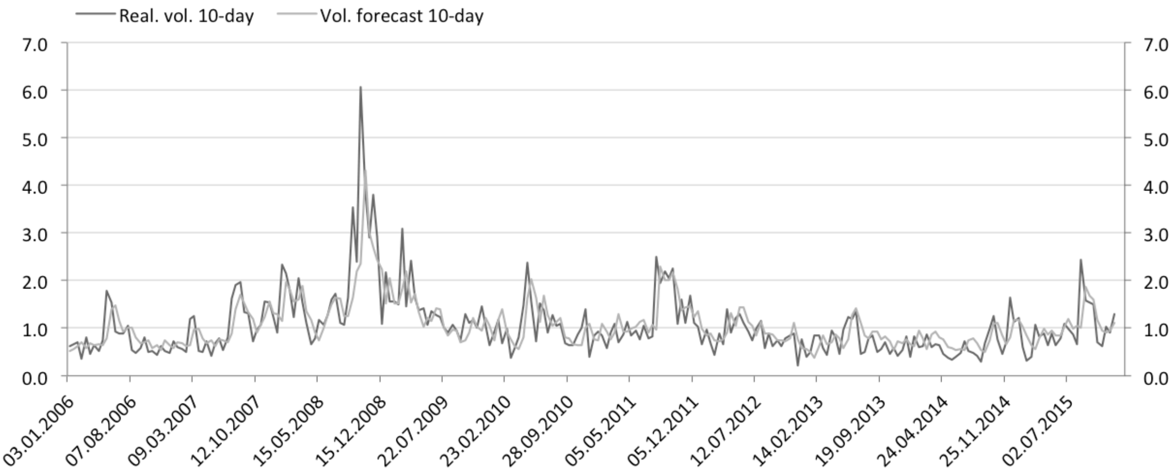


Fig. 7.2.3: Normal AR(1) GJR-GARCH(1,1), 20-day horizon realized (proxy) and forecasted volatility

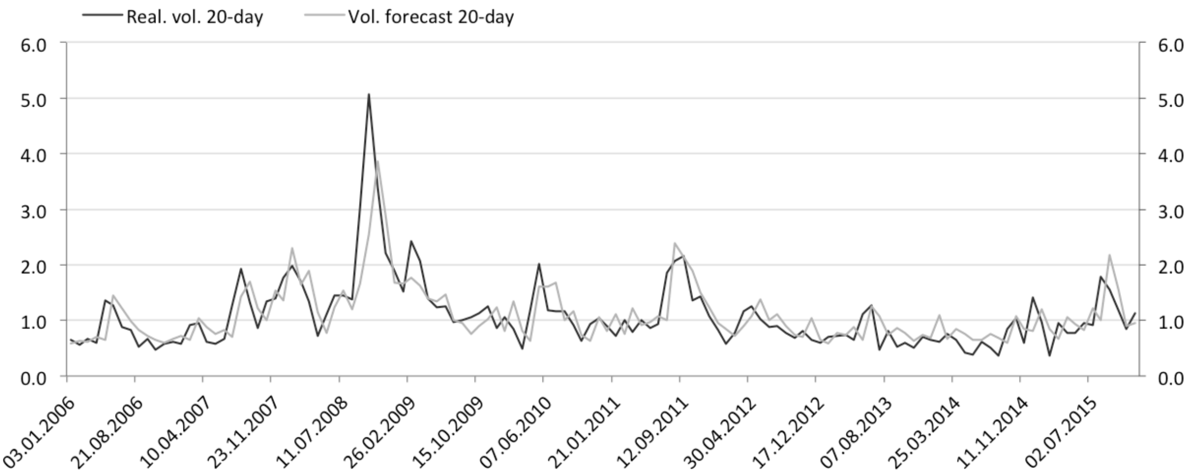


Fig. 7.2.4 illustrates how conditional variance parameter values changes over time for a normal GJR-GARCH model with a first order autoregressive mean. The opposite trends in beta and delta are evidence of declining volatility persistence. The symmetric shock parameter has little significance from the start. In late 1998 alpha jumps and then collapses to zero. This is when the rolling window begins to include data from the market crash of 2008.

Returns appear to be autocorrelated, although not through the whole sample period. Standard GARCH models perform better with an autoregressive conditional mean. For GJR-GARCH models this depends on which loss function is looked at. The significance of the first order lag in the above-mentioned model changes considerably during the estimation procedure. This is illustrated in figure 7.2.5.

Fig. 7.2.4: Normal AR(1) GJR-GARCH(1,1), model coefficients over time
The horizontal axis shows the estimation window starting date

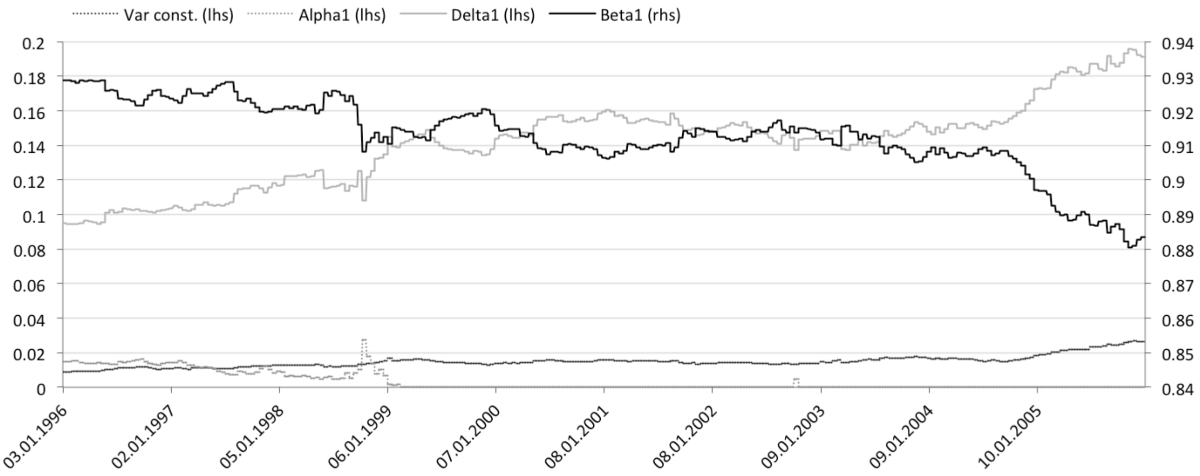
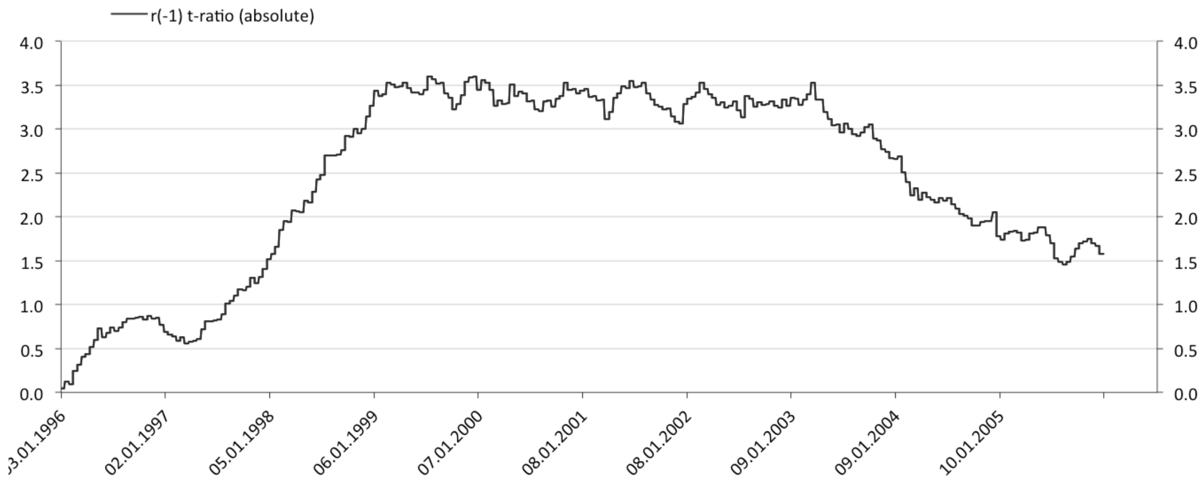


Fig. 7.2.5: Normal AR(1) GJR-GARCH(1,1), $\hat{\phi}$ absolute t-ratio
The horizontal axis shows the estimation window starting date



The SMA has the lowest historical accuracy. Switching to exponential weighting gives a notable reduction in loss function values. Of the two presented, the one with a lambda value of 0.95 is the most precise.

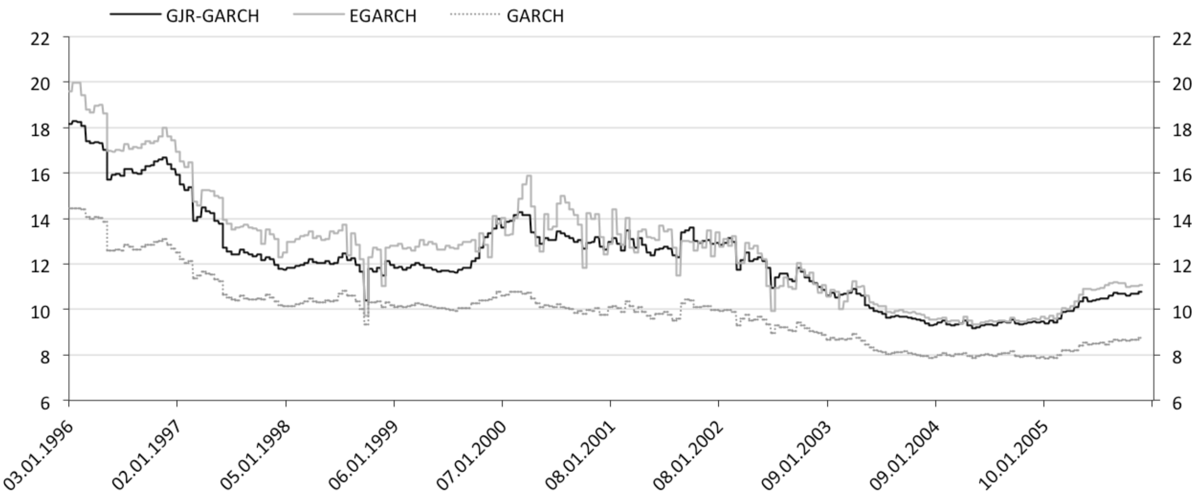
7.3 OMXS30

Table 7.3.1: OMXS30 out-of-sample forecasting results

	10 days		20 days	
	MSE	QLIKE	MSE	QLIKE
SMA	3.7891	1.5720	3.9426	1.6205
EWMA 0.92	3.1233	1.5250	2.9578	1.6018
EWMA 0.95	3.4810	1.5233	3.2553	1.5938
Const. in cond. mean				
GARCH normal	2.9787	1.4852	2.7697	1.5330
GARCH symmetric t	3.1762	1.4902	2.9902	1.5401
GARCH skewed t	3.1610	1.4905	2.9591	1.5403
GJR-GARCH normal	2.6233	1.4745	2.6353	1.5333
GJR-GARCH symmetric t	2.7022	1.4775	2.7177	1.5386
GJR-GARCH skewed t	2.7334	1.4784	2.8065	1.5413
EGARCH normal	3.1536	1.4905	3.2670	1.5528
EGARCH symmetric t	3.0621	1.4937	3.1469	1.5575
EGARCH skewed t	3.2300	1.4990	3.5284	1.5815
Const. + lag in cond. mean				
GARCH normal	3.0157	1.4847	2.7473	1.5322
GARCH symmetric t	3.1768	1.4897	2.9859	1.5395
GARCH skewed t	3.1610	1.4905	2.9799	1.5391
GJR-GARCH normal	2.6501	1.4745	2.6409	1.5333
GJR-GARCH symmetric t	2.7018	1.4777	2.7269	1.5388
GJR-GARCH skewed t	2.7256	1.4784	2.8118	1.5408
EGARCH normal	3.1373	1.4897	3.2566	1.5515
EGARCH symmetric t	2.5983	1.4816	3.1483	1.5569
EGARCH skewed t	3.2322	1.4988	3.5258	1.5820

The OMXS30 unconditional return distribution is heavy-tailed. This is also the case for fitted Student t conditional distributions, as can be seen in figure 7.3.1. Degree of freedom parameter values are low and stable throughout the whole estimation period. Even so, the lowest loss function values for GJR and standard GARCH models are obtained with a normal error distribution. Choosing the skewed t distribution typically results in the highest values for all three model types.

Fig. 7.3.1: Symmetric t GARCH models, degree of freedom parameter value over time
The horizontal axis shows the estimation window starting point



Asymmetric models easily outperform standard GARCH models when the forecast horizon is 10 days. When extending it to 20 days the difference shrinks significantly. GJR-GARCH models have consistently low values. On the other hand, for some exponential GARCH models, values are very high, especially with a 20-day forecast period. The normal GJR-GARCH model has the lowest loss function values in two of four columns. Taking everything into account, it should be the preferred model. In figure 7.3.2 volatility forecasts generated by this model is compared with estimated volatility on a 10-day horizon. In the two remaining columns other model types have it beat by a small margin. One of these is the symmetric t EGARCH model. Its accuracy increases dramatically on a 10-day forecast horizon with a first order autoregressive conditional mean. This puts it slightly ahead of the normal GJR-GARCH model when using the MSE loss function for evaluation. It falls out of favor when increasing the forecast length.

Specifying the conditional mean as autoregressive does not generally result in markedly lower loss function values. The fact that they are not consistently different would suggest that autocorrelation is not completely absent from sample returns.

Fig. 7.3.2: Normal GJR-GARCH(1,1), 10-day horizon realized (proxy) and forecasted volatility

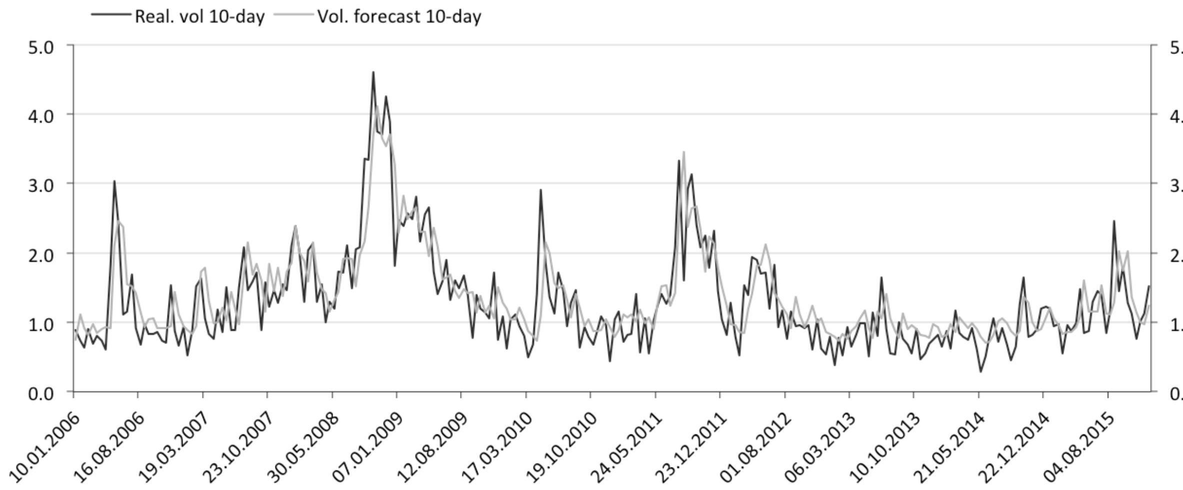


Fig. 7.3.3: Normal GJR-GARCH(1,1), model coefficients over time

Horizontal axis shows the estimation window starting date

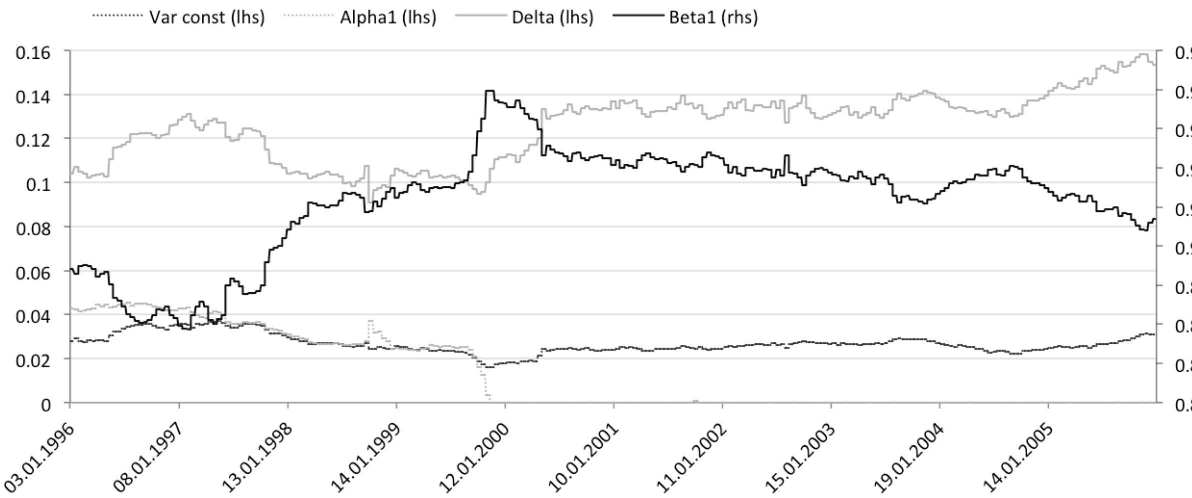


Figure 7.3.3 illustrates how conditional variance parameter values in the preferred model changes over time. The level of persistence in conditional variance, represented by beta, fluctuates quite a bit during the sample period with no visible trend. Alpha reaches zero when the rolling window starts to include data for late 2009. From this point the leverage parameter is the only shock parameter.

Naïve models have the highest loss function values, with a few exceptions. If limited to these, the SMA should be avoided in favor of EWMA alternatives. MSE values belonging to the latter are similar to those of some exponential GARCH models.

7.4 Summary

There are many similarities between the European equity indices that this study is based on, which inevitably extend to forecasting results. Samples for all three indices have unconditional return distributions that exhibit high excess kurtosis. Assumed conditional error distributions for Student *t* models also have this characteristic. The highest log-likelihood values are obtained with low degree of freedom parameter values. Out-of-sample forecasting results contradict these observations. Staying with the normal distribution generally leads to more accurate forecasts. For GJR-GARCH models loss function values are consistently low in a relative sense. Especially normal GJR-GARCH does well. It is robust and it accommodates the leverage effect. For exponential GARCH models, on the other hand, results have a much higher dispersion. The relative performance of EGARCH models seems to deteriorate rapidly as the forecast horizon is increased.

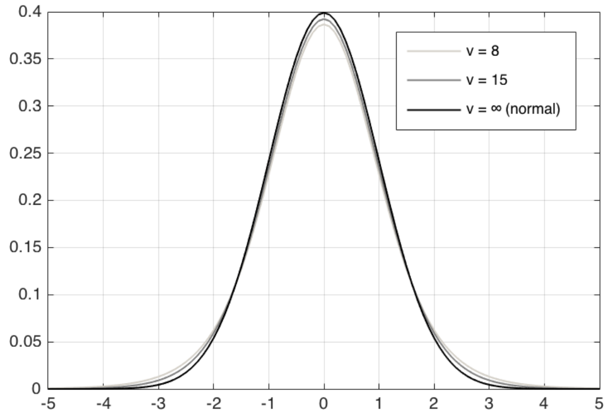


Fig. 7.4.1: Student's *t* PDF examples

Specifying the mean equation as autoregressive results in consistently lower loss function values for standard GARCH models across all three indices. For GJR and exponential GARCH models it's harder to discern a pattern.

Naïve models are easier to apply since they do not require regression software and no formulas are needed to generate forecasts. To at least have a forecast is preferable to not having any. They are however easily outperformed by most of the GARCH models.

The evolution in GJR-GARCH model coefficients shows that the three indices have an important feature in common. The leverage parameter gradually gains influence relative to other parameters, as more and more post-crisis data is included in the rolling estimation window. In other words, market nervousness increases during the sample period. It would be interesting to see how this development relates to history in a much larger sample.

Enlarged charts of estimated volatility and forecasted volatility by the best performing models can be found in Appendix C.

8 Conclusions

The main objective in this study has been to determine which univariate GARCH specification and conditional distribution is best suited for variance forecasting, judging from historical performance. The analysis is based on data samples for DAX30, FTSE100 and OMXS30, but the main conclusions are also relevant for equity indices in general. Results are for the most part in line with earlier comparable research. Information criteria provide an inaccurate indication of which model should be preferred for forecasting. When the focus is on ensuring an adequate fit the model often becomes overfitted.

MSE and QLIKE treat outliers very differently. This result in some dissimilarities, but generally they agree on which models should be preferred. Because of the leverage effect, GJR- and exponential models can be expected to outperform. Results confirm this, with some exceptions. The GJR specification is a safe choice for all three indices, especially if combined with a normal error distribution. However, even though GJR-GARCH should be preferred, the fact that parameter constraints are necessary would suggest that this specification is not optimal for the sample data. EGARCH results are more ambiguous. With a 10-day forecast horizon normal or symmetric *t* EGARCH should be the second choice. If the horizon is extended to 20 days even standard GARCH models do better. Skewed *t* EGARCH should be avoided altogether. The fact that this model often has the best goodness of fit confirms that it is overfitted.

The relative performance of non-normal models is surprising. All three indices have non-zero skewness and excess kurtosis. It is understandable that skewed *t* models do not outperform. The skewness is simply not pronounced enough to warrant optimization of it. On the other hand, levels of kurtosis are very high. This is also reflected by fitted Student *t* conditional distributions. The reason behind the disappointing results of symmetric *t* models may be that the rest of the density is given less attention. A distribution that could also accommodate greater 'peakedness' would probably do more for forecasting accuracy.

If limited to naïve models the EWMA should always be preferred, at least for the forecast horizons considered here. Though, results provide no solid indication of which lambda coefficient is optimal. Just by upgrading to the most basic GARCH model loss function values drop meaningfully.

A number of extensions of this research are attractive. By using intra-day data to estimate variance results would become clearer. Measuring forecasting performance over a longer sample period could also help in this regard.

Other loss functions could also be considered. For example, for a market trading strategy the relevance of the study would increase by including financial loss functions.

The extension that should be given the most attention relates to non-normality. There are few prior GARCH studies in which skewed *t* distributions have been considered. It would therefore be of interest how skewed *t* models perform with sample data that exhibits more pronounced asymmetry than the samples included here.

References

Alexander, C. (2008): *Market Risk Analysis Vol. II: Practical Financial Econometrics* John Wiley & Sons Ltd., West Sussex

Andersen, T.G. and T. Bollerslev (1998): "Answering the sceptics: Yes, standard volatility models do provide accurate forecasts*", International Economic Review vol. 39*, No. 4, 885-905

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

Bollerslev, T. (1987): "A conditionally heteroskedastic time series model for speculative prices and rates of return", *The Review of Economics and Statistics 69*, No. 3, 542-547

Bollerslev, T., R.F. Engle and D.B. Nelson (1994): "ARCH models", *Handbook of Econometrics 4*, 2960-3038

Black, F. and M. Scholes (1973): "The pricing of options and corporate liabilities", *Journal of Political Economy 81*, No. 3, 637-654

Black, F. (1976): "Studies in stock price volatility changes. Proceedings of the 1976 business meeting of the business and economics section", *American Statistical Association*, 177-181

Cont, R. (2005): "Volatility Clustering in Financial Markets: Empirical Facts and Agent Based Models", Ecole Polytechnique, University of Paris-Saclay

Dijk, P.H. and D. Franses (1996): "Forecasting stock market volatility using (non-linear) GARCH models", *Journal of Forecasting 15*, 229-235

Ding, Z., R. Engle and C.W.J. Granger (1993): "A long memory property of stock market returns and a new model", *Journal of Empirical finance 1*, 83-106

Engle, R.F. (1982): "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation", *Econometrica 50*, No. 4, 987-1007

Engle, R.F. and V.K. Ng (1993): "Measuring and testing the impact of news on volatility", *The Journal of Finance 48*, No. 5, 1749-1778

Fama, E. (1965): “The Behavior of Stock-Market Prices”, *The Journal of Business 38*, No. 1, 34-105

Glosten, L.R, R. Jagannathan and D.E. Runkle (1993): “On the relation between the expected value and the volatility of the nominal excess return on stocks”, *Journal of Finance 48*, No. 5, 1779-1801

Lambert and Laurent (2001): “Modelling skewness dynamics using GARCH-type models and a skewed Student density”, Mimeo, University of Liège

Mandelbrot, B. (1963): “The Variation of Certain Speculative Prices”, *The Journal of Business 36*, No.4, 394-419

Nelson, D.B. (1991): “Conditional heteroscedasticity in asset returns: A new approach”, *Econometrica 59*, No.4, 347-370

Parkinson, M. (1980): “The extreme value method for estimating the variance of the rate of return”, *The Journal of Business 53*, No.1, 61-65

Patton, A.J. (2011): “Volatility forecast comparison using imperfect volatility proxies”, *Journal of Econometrics 160*, No.1, 246-256

Patton, A.J. and K. Sheppard (2009): “Evaluating volatility and correlation forecasts”, In: Andersen, T.G., R.A. Davis, J.P. Kreiss and T. Mikosch (Eds.), *The Handbook of Financial Time Series*, Springer Verlag, Berlin

Peters (2001): “Estimating and forecasting volatility of stock indices using asymmetric GARCH models and (skewed) Student-t densities”, EAA Business School, University of Liège

Taylor, S.J. (1986): “Forecasting the volatility of currency exchange rates”, *International Journal of Forecasting 3*, 159-170

Appendix A: Formulae for statistics and tests

Jarque-Bera normality test

The distribution of a normally distributed random variable is characterized by the first two moments, that is, the mean and the variance. The third and fourth moment are known as its skewness and kurtosis, respectively. In a normal distribution the coefficients of skewness and kurtosis are jointly zero, which is what the test checks for

$$\mathrm{JB} = \mathrm{T}\left[\frac{\mathrm{b}_1^2}{6} + \frac{(\mathrm{b}_2 - 3)^2}{24}\right]$$

where b_1 and b_2 denote the coefficients of sample skewness and sample kurtosis, respectively. T is the sample size. The test statistic asymptotically follows a $\chi^2(2)$ under the null hypothesis that the distribution of the series is symmetric and mesokurtic.

Ljung-Box Q-test

Autocorrelation coefficients can vary significantly between lags. The Ljung-Box Q-test is a quantitative alternative that checks for autocorrelation at multiple lags jointly

$$\mathrm{Q(m)} = \mathrm{T(T + 2)} \int_{\mathrm{k}=1}^{\mathrm{m}} \frac{\hat{\tau}_{\mathrm{k}}^2}{\mathrm{T} - \mathrm{k}}$$

where $\hat{\tau}_k$denotes the estimated autocorrelation coefficient at lag k, m is the maximum lag length and T is the sample size. The Q-statistic is asymptotically distributed as a $\chi^2(m)$ under the null hypothesis that all m autocorrelation coefficients are zero.

Engle's ARCH LM test

The ARCH test, proposed by Engle (1982), checks for the presence of autoregressive conditional heteroscedastic effects by measuring autocorrelation in the squared series. When used to assess whether the conditional variance of a time series is suitable for ARCH modeling the test is normally conducted on the mean equation innovation ε_t, which is given by

$$\varepsilon_{\mathrm{t}} = \mathrm{r}_{\mathrm{t}} - \hat{\mu}_{\mathrm{t}}$$

where $\hat{\mu}_t$ is the conditional mean and r_t is the observed return. Autoregressive lags should be included if returns are autocorrelated. The test regression can be written as

$$\varepsilon_{\mathrm{t}}^2 = \gamma_0 + \gamma_1\varepsilon_{\mathrm{t}-1}^2 + \gamma_2\varepsilon_{\mathrm{t}-2}^2 + \cdots + \gamma_{\mathrm{m}}\varepsilon_{\mathrm{t}-\mathrm{m}}^2 + \mathrm{u}_{\mathrm{t}}$$

where u_t is an error term and m is the number of lags. The test statistic is defined as TR^2, where T is the sample size and R^2 represents the coefficient of determination. It is distributed as a $\chi^2(m)$ under the null hypothesis that all m coefficients are zero.

Engle and Ng's sign bias test

The sign bias test, as defined by Engle and Ng (1993), indicates whether positive and negative shocks of equal size have a different effect on future conditional variance. The test is conducted on the standardized residuals of a standard GARCH model, defined as $z_t = \varepsilon_t/\sqrt{h_t}$. The regression for the sign bias test is then given by

$$z_t^2 = \theta_0 + \theta_1 S_{t-1}^{-} + u_t$$

where u_t is the error term and S_{t-1}^{-} is an indicator dummy that takes the value 1 if $\varepsilon_{t-1} < 0$ and zero otherwise. If positive and negative shocks to ε_{t-1} impact differently upon the conditional variance, then θ_1 will have a statistically significant t-statistic.

Appendix B: Log-likelihood functions

If innovations are drawn from a normal distribution the log-likelihood function is defined by

$$\ln L(\theta) = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log(h_t) - \frac{1}{2}\sum_{t=1}^{T}\frac{\varepsilon_t^2}{h_t}$$

where ε_t represents innovations and h_t the conditional variance. If innovations instead are drawn from a symmetric Student *t* distribution[1] the log-likelihood function is defined by

$$\ln L(\theta) = T\left(\ln\Gamma\left(\frac{v+1}{2}\right) - \ln\Gamma\left(\frac{v}{2}\right) - \frac{1}{2}\ln[\pi(v-2)]\right) - \frac{1}{2}\sum_{t=1}^{T}\left[\ln(h_t) + (1+v)\ln\left(1+\frac{z_t^2}{v-2}\right)\right]$$

where z_t is given by $\varepsilon_t/\sqrt{h_t}$, v is the degree of freedom of the distribution, and $\Gamma(\cdot)$ is the gamma function. Lastly, the log-likelihood function when innovations are assumed to follow a skewed Student *t* distribution[2]

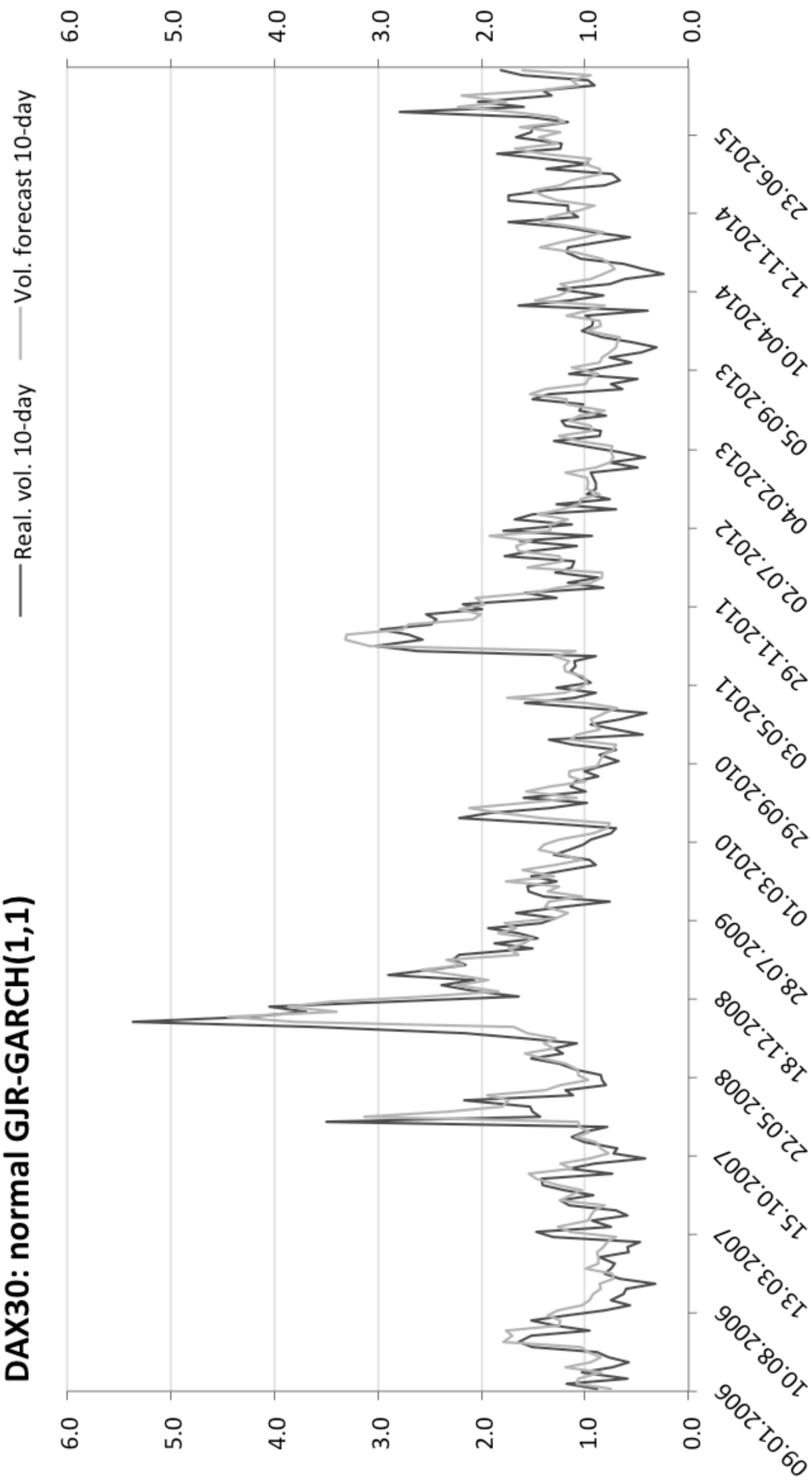
$$\ln L(\theta) = T\left(\ln\Gamma\left(\frac{v+1}{2}\right) - \ln\left(\frac{v}{2}\right) - \frac{1}{2}\ln\big(\pi(v-2)\big) + \ln\left(\frac{2}{\xi+(1/\xi)}\right) + \ln(s)\right) - \frac{1}{2}\sum_{t=1}^{T}\left((\ln(h_t) + (1+v)\ln\left(1+\frac{(sz_t+m)^2}{v-2}\xi^{-2I_t}\right)\right)$$

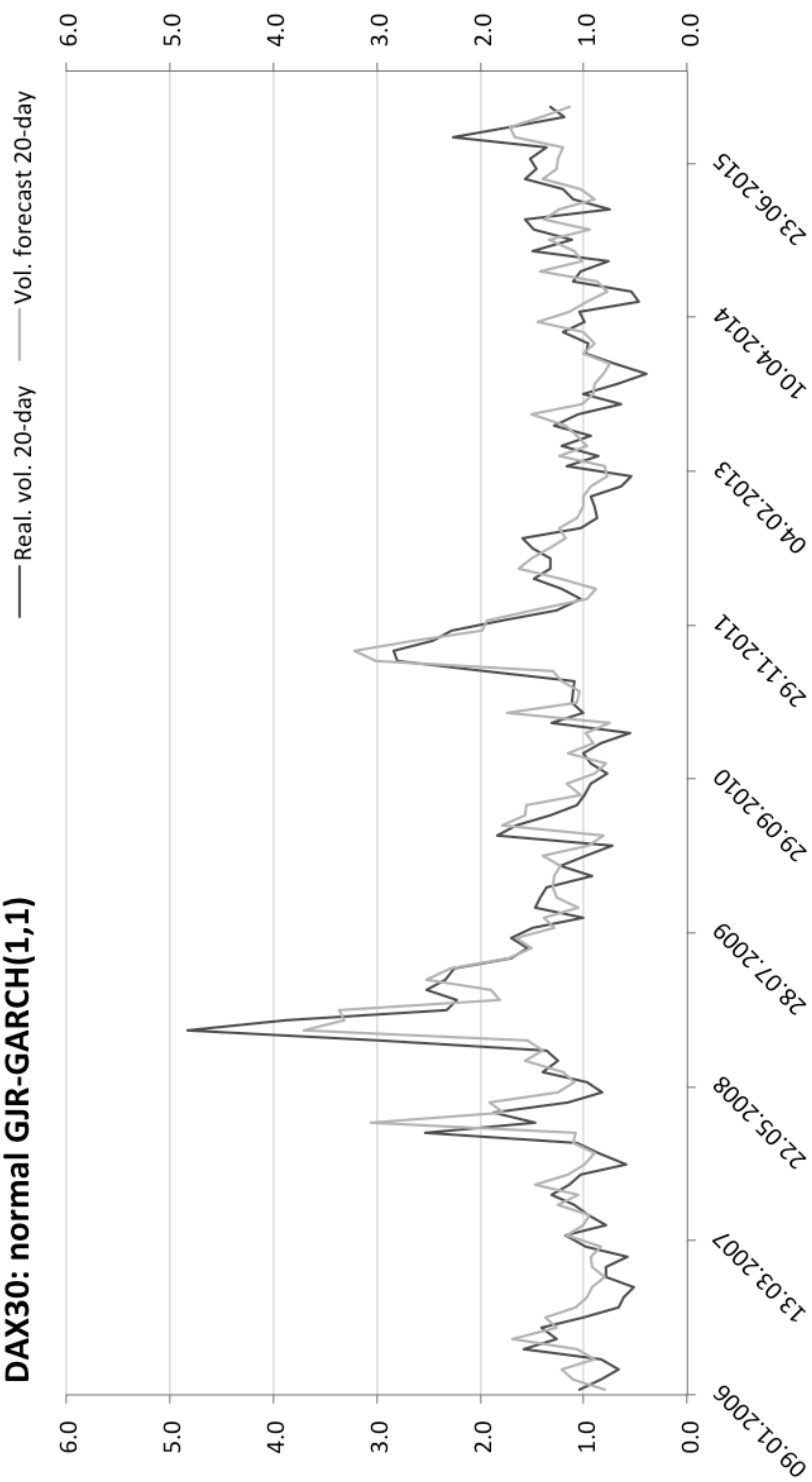
where ξ is the asymmetry parameter, and

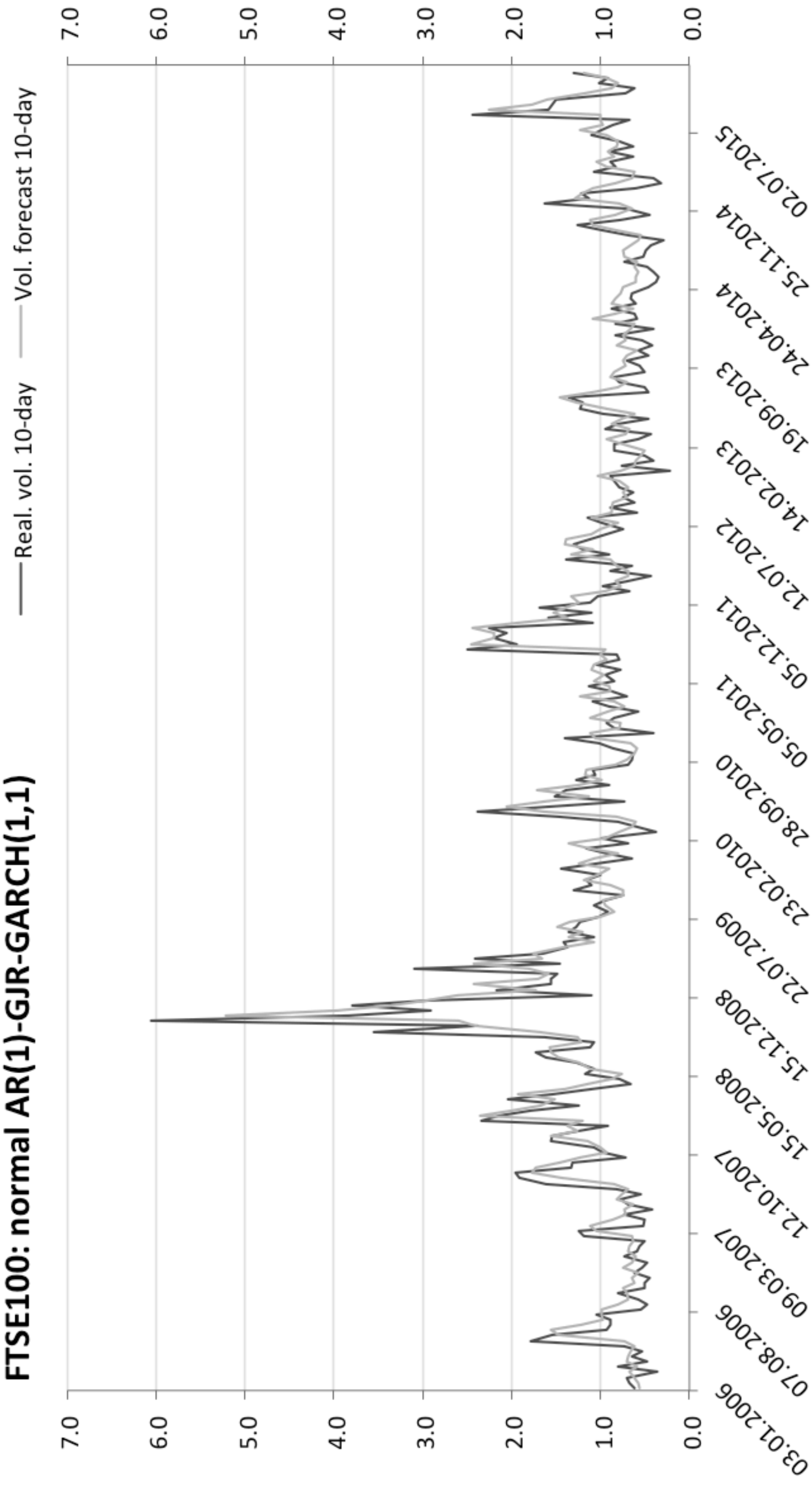
$$I_t = \begin{cases} 1 \text{ if } z_t \geq -\frac{m}{s} \\ -1 \text{ if } z_t < -\frac{m}{s} \end{cases}, m = \frac{\Gamma\left(v+\frac{1}{2}\right)\sqrt{v-2}}{\sqrt{\pi}\Gamma\left(\frac{v}{2}\right)}\left(\xi - \frac{1}{\xi}\right) \text{ and } s = \sqrt{\left(\xi^2 + \frac{1}{\xi^2} - 1\right) - m^2}$$

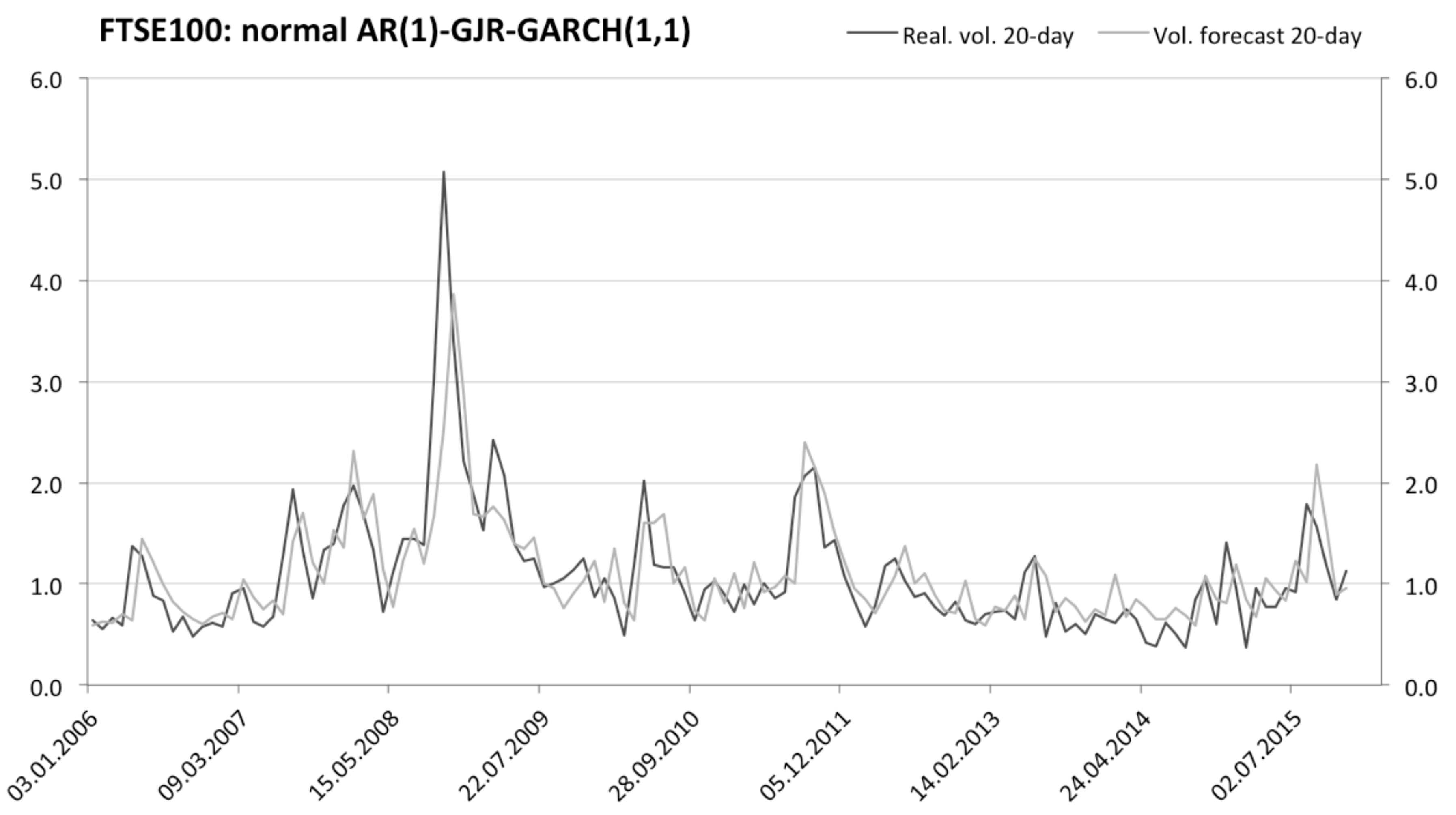
[1,2] extended to the GARCH framework by Bollerslev (1987), and Lambert and Laurent (2001), respectively.

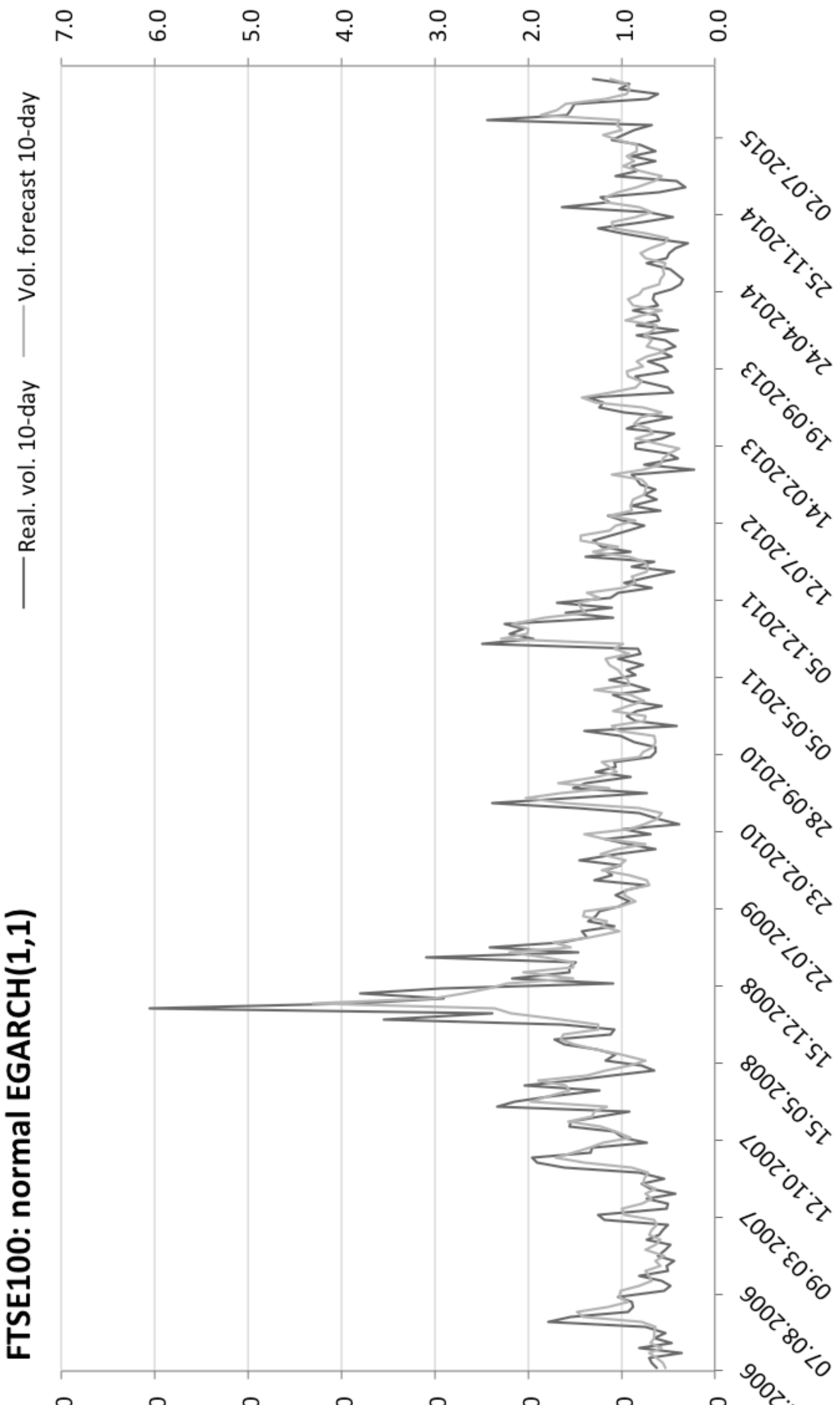
Appendix C: Volatility charts

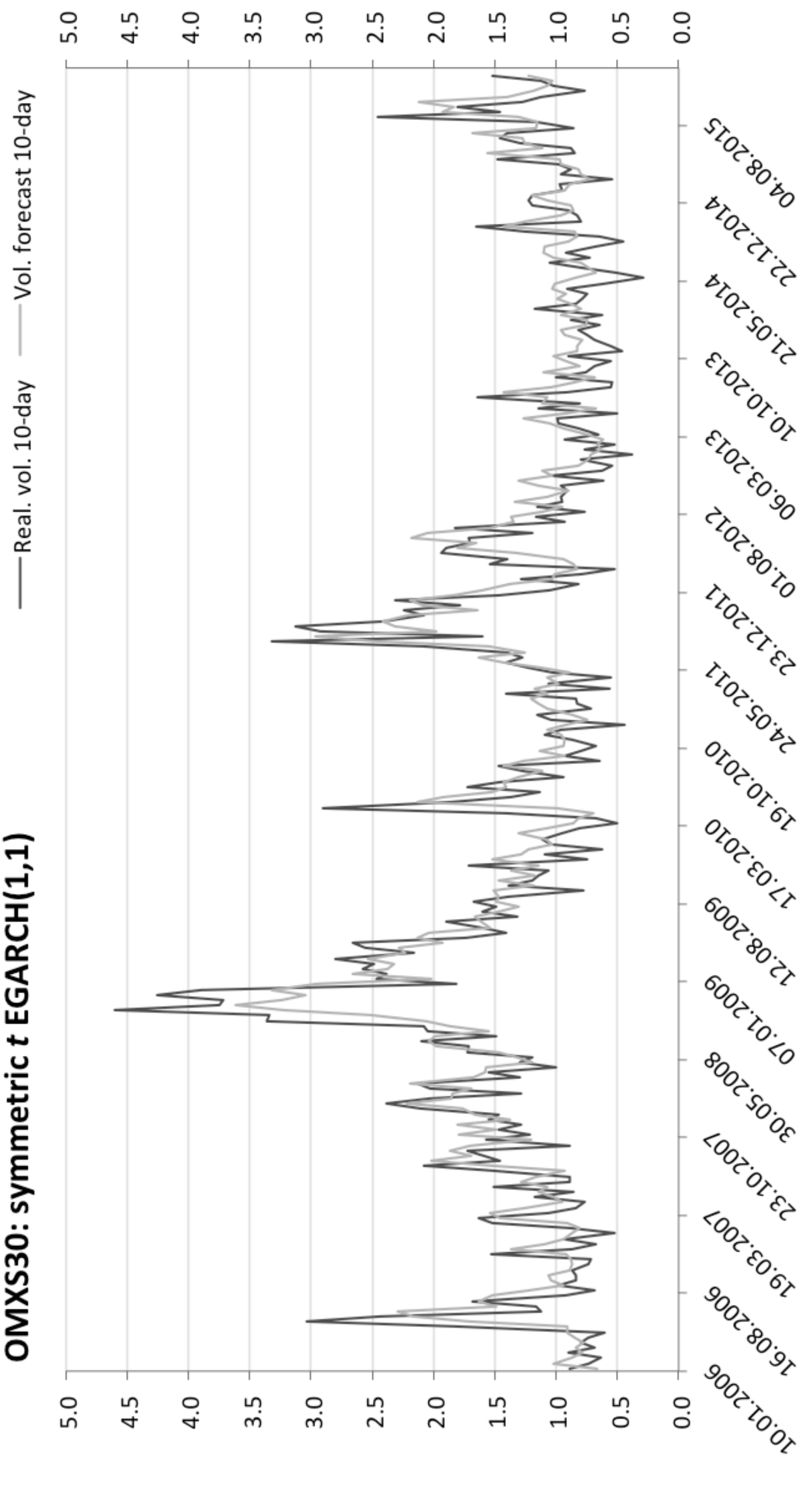



DAX30: normal GJR-GARCH(1,1)
Real. vol. 20-day
Vol. forecast 20-day
0.0
1.0
2.0
3.0
4.0
5.0
6.0
09.01.2006
13.03.2007
22.05.2008
28.07.2009
29.09.2010
29.11.2011
04.02.2013
10.04.2014
23.06.2015

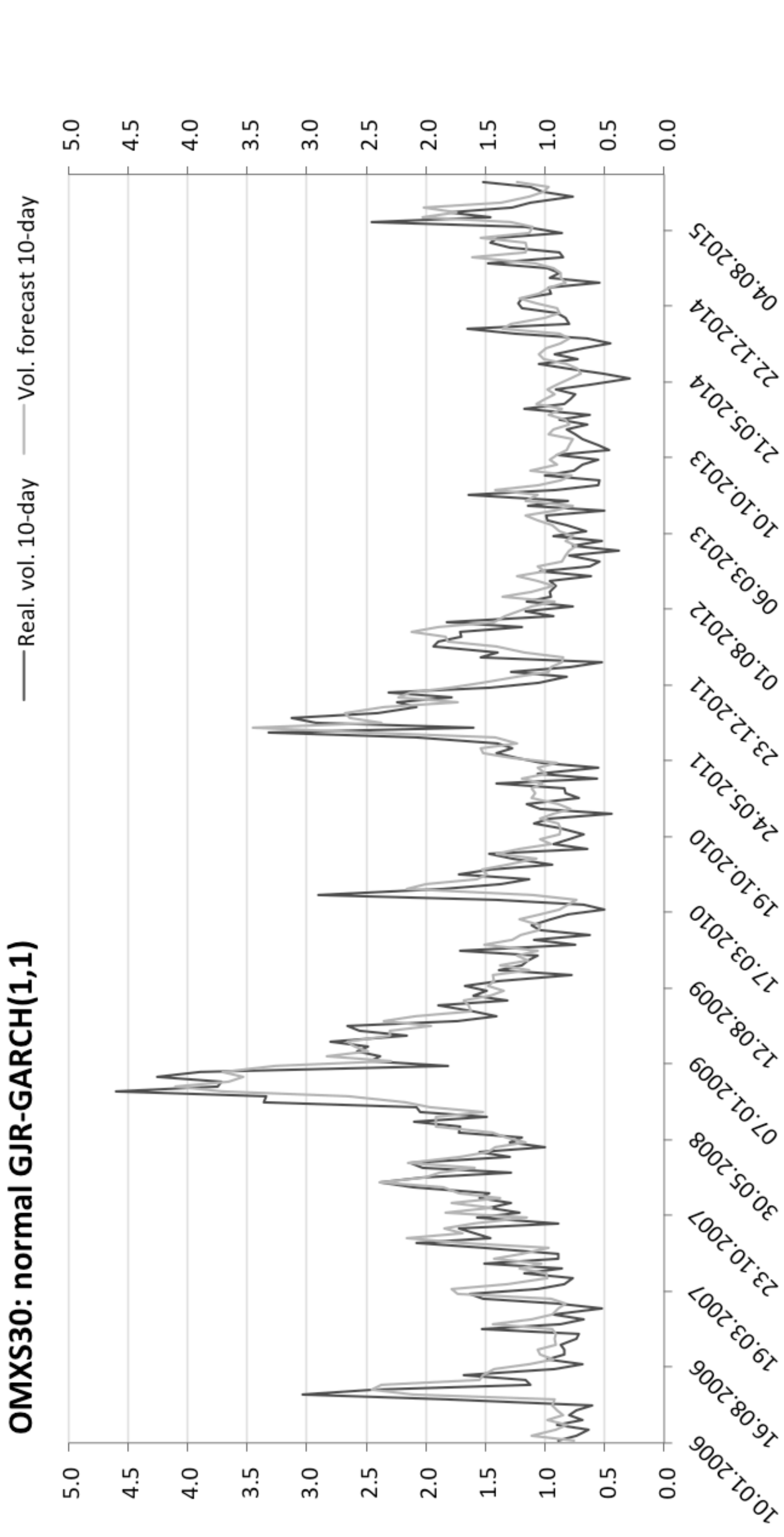

FTSE100: normal AR(1)-GJR-GARCH(1,1)
Real. vol. 10-day
Vol. forecast 10-day
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
03.01.2006
07.08.2006
09.03.2007
12.10.2007
15.05.2008
15.12.2008
22.07.2009
23.02.2010
28.09.2010
05.05.2011
05.12.2011
12.07.2012
14.02.2013
19.09.2013
24.04.2014
25.11.2014
02.07.2015


FTSE100: normal AR(1)-GJR-GARCH(1,1)
Real. vol. 20-day
Vol. forecast 20-day
6.0
5.0
4.0
3.0
2.0
1.0
0.0
03.01.2006
09.03.2007
15.05.2008
22.07.2009
28.09.2010
05.12.2011
14.02.2013
24.04.2014
02.07.2015

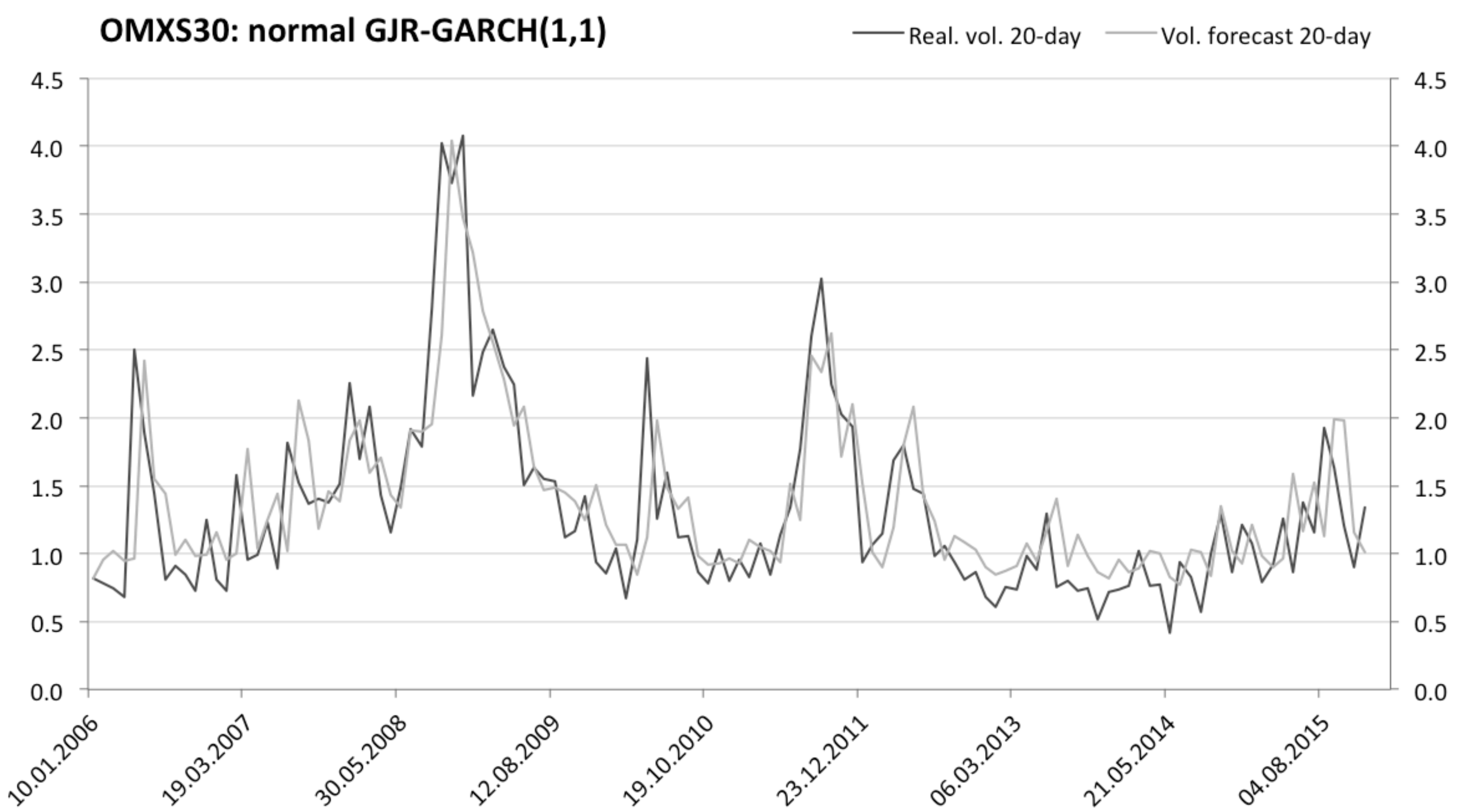

FTSE100: normal EGARCH(1,1)
Real. vol. 10-day
Vol. forecast 10-day
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
03.01.2006
07.08.2006
09.03.2007
12.10.2007
15.05.2008
15.12.2008
22.07.2009
23.02.2010
28.09.2010
05.05.2011
05.12.2011
12.07.2012
14.02.2013
19.09.2013
24.04.2014
25.11.2014
02.07.2015

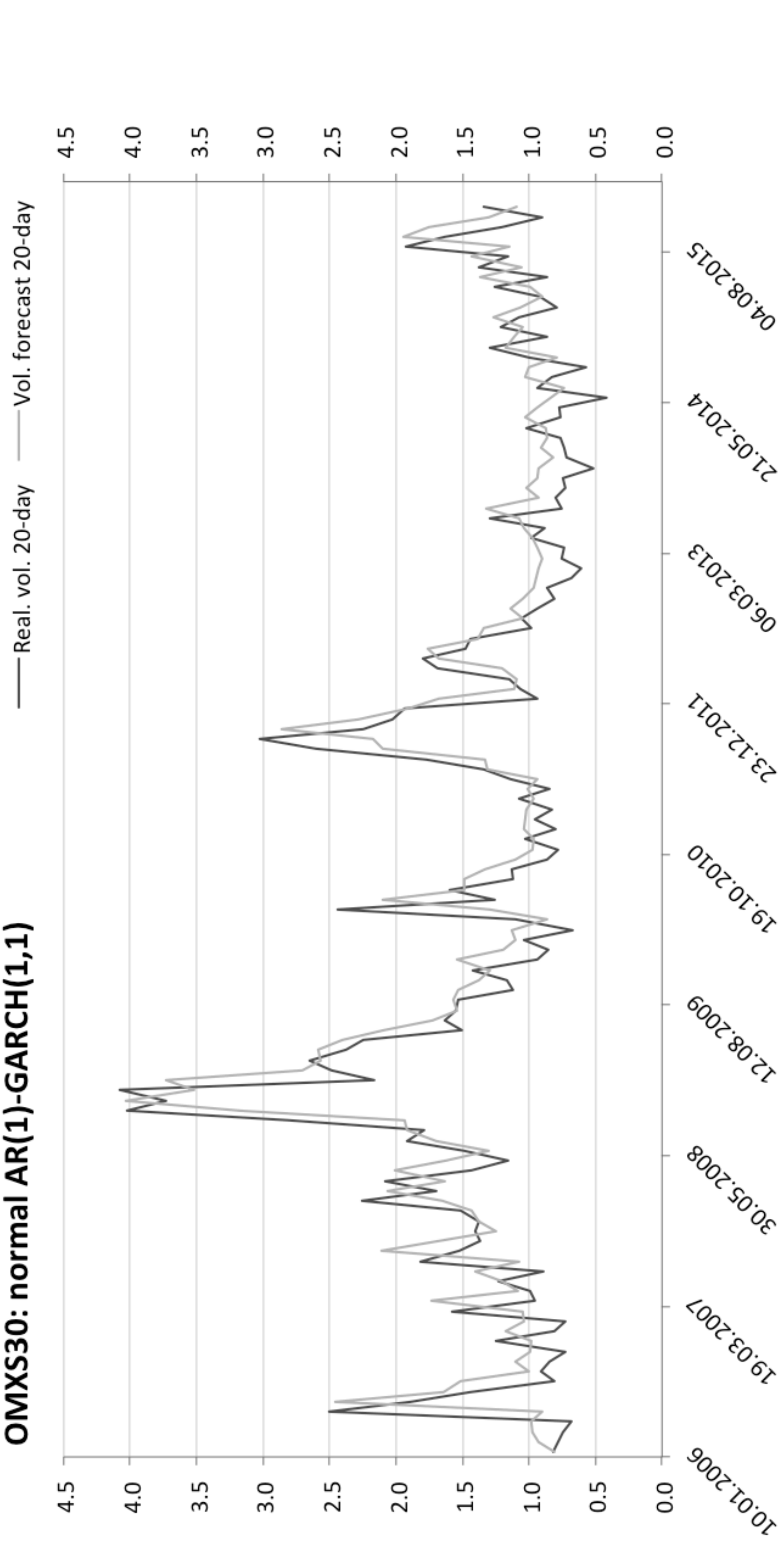


OMXS30: symmetric t EGARCH(1,1)
Real. vol. 10-day
Vol. forecast 10-day
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
10.01.2006
16.08.2006
19.03.2007
23.10.2007
30.05.2008
07.01.2009
12.08.2009
17.03.2010
19.10.2010
24.05.2011
23.12.2011
01.08.2012
06.03.2013
10.10.2013
21.05.2014
22.12.2014
04.08.2015



OMXS30: normal GJR-GARCH(1,1)
Real. vol. 10-day
Vol. forecast 10-day
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
10.01.2006
16.08.2006
19.03.2007
23.10.2007
30.05.2008
07.01.2009
12.08.2009
17.03.2010
19.10.2010
24.05.2011
23.12.2011
01.08.2012
06.03.2013
10.10.2013
21.05.2014
22.12.2014
04.08.2015


OMXS30: normal GJR-GARCH(1,1)
Real. vol. 20-day
Vol. forecast 20-day
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
10.01.2006
19.03.2007
30.05.2008
12.08.2009
19.10.2010
23.12.2011
06.03.2013
21.05.2014
04.08.2015


OMXS30: normal AR(1)-GARCH(1,1)
Real. vol. 20-day
Vol. forecast 20-day
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
10.01.2006
19.03.2007
30.05.2008
12.08.2009
19.10.2010
23.12.2011
06.03.2013
21.05.2014
04.08.2015