Assignment 4 - Econometrics 2

Department of Economics - University of Copenhagen

Sunday, 06.12.2020, 23:00

1 Introduction

Returns of financial assets tend to be very volatile. Their volatility often appears to be somewhat clustered, i.e. during some time spans the returns' variation is very high and during others rather low. According to some behavioral theories, these movements in returns and volatility can be explained by unexpected movements in trading volumes (so called surprise trading volume), instead of the overall trading volume per se. In this assignment we are asked to analyze if the surprise trading volume of the same day affects the conditional mean and the conditional variance of the German stock index. The best approach to answer this question is to build a model for the daily change in the stock market index using the family of (G)ARCH-models as they are ideal to model volatility clustering. We end up with an EGARCH(5,4) as our preferred model and find huge effects of surprise trading volume on mean and variance of DAX returns.

2 Description of Data

For our analysis we use daily data[1] of the German blue chip stock market index, DAX 30, which includes the 30 largest German companies and represents about 80% of total market capitalisation of all listed limited-liability firms[2]. The considered time period covers 17 September 2001 to 30 October 2020, accounting for a total number of 4854 observations. Besides DAX, the data set contains a variable measuring the total Trading Volume (Vol) in number of stocks traded. Furthermore dummies for the day before and after a closing day (FirstTrDay / LastTrDay) as well as monthly dummies (isMonth) are present. We use the variable transformations shown in Equation 1. By taking first differences for the $log(DAX)$ we obtain a stationary time series, which shows no autocorrelation (Figure 1). The 50-day moving average for the trading volume ($logVol50$) is used later as the expected trading volume.

$$\begin{aligned} DAX &\longrightarrow log(DAX) \longrightarrow \textbf{DlogDAX: } 100 \cdot (logDAX_t - logDAX_{t-1}) \\ Vol &\longrightarrow log(Vol) \longrightarrow \textbf{logVolMA50: } logVol50 \end{aligned} \tag{1}$$

3 Economic Theory

Agents in the financial markets are interested in the mean and the variance of the assets they are holding. A special characteristic of financial data in time series is its non-constant volatility. Looking at Figure 1 we clearly see periods of volatility clustering throughout time. In order to model this appropriately, we consider the generalized autoregressive conditional heteroscedasticity (GARCH) model.

3.1 GARCH processes

Recalling simple dynamic autoregressive models (AR), past events were used to estimate the conditional mean of a variable today. This usage of a conditional mean is now extended to the conditional variance as well. A GARCH(1,1) model is characterized as follows:

$$\begin{aligned} y_t &= \mu_t + \epsilon_t \\ \epsilon_t &= \sigma_t z_t & z_t|I_{t-1} \sim iid(0,1) \\ \sigma_t^2 &= \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 & \omega > 0, \alpha \geq 0, \beta \geq 0 \end{aligned} \tag{2}$$

The variation of ϵ_t is decomposed into two terms, z_t assumed to be independent and identically distributed (iid) and the conditional standard deviation σ_t. The shape of the distribution is not further specified and can be chosen or estimated in order to fit the data as good as possible.

[1] The data is downloaded from https://stooq.com/.

[2] Source:https://web.archive.org/web/20171204061235/http://deutsche-boerse.com/dbg-de/presse/pressemitteilungen/Deutsche-Boerse-startet-neuen-XDAXDAX-Index/2386400

The conditional variance of the GARCH process is dependent on the information set of past events I_{t-1} and defined as $\sigma_t^2 = E(\epsilon_t^2|I_{t-1})$. The variance in Equation 2 can be further decomposed, where v_t is the uncorrelated surprise to the conditional squared residual ϵ_t^2.

$$\begin{aligned}
\epsilon_t^2 &= E(\epsilon_t^2|I_{t-1}) + v_t \qquad E(v_t|I_{t-1}) = 0 \\
\sigma_t^2 &= \epsilon_t^2 - v_t \\
\sigma_t^2 &= \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \\
\epsilon_t^2 - v_t &= \omega + \alpha\epsilon_{t-1}^2 + \beta(\epsilon_{t-1}^2 - v_{t-1}) \\
\epsilon_t^2 &= \omega + \underbrace{(\alpha+\beta)\epsilon_{t-1}^2}_{AR(1)\ term} + \underbrace{v_t - \beta v_{t-1}}_{MA(1)\ term}
\end{aligned} \tag{3}$$

The last row of Equation 3 shows how the residuals of the initial model can be reformulated to a ARMA(1,1) model for the squared innovation. In order for ϵ_t to be stationary and stable $\alpha + \beta < 1$. The necessity of this assumption is also clear when we derive the unconditional variance (Equation 4), by taking the unconditional expected value of the squared residuals, with $\epsilon_t = \epsilon_{t-1}$ due to stationarity:

$$\begin{aligned}
\text{LIE:} \qquad & E(v_t|I_{t-1}) = E(E(v_t|I_{t-1})) \\
E(\epsilon_t^2) &= \omega + (\alpha+\beta)E(\epsilon_t^2) + E(v_t) - \beta E(v_{t-1}) \\
E(\epsilon_t^2) &= \omega + (\alpha+\beta)E(\epsilon_t^2) + E(E(v_t|I_{t-1})) - \beta E(E(v_{t-1}|I_{t-1})) \\
\underbrace{E(\epsilon_t^2)}_{\sigma_t^2}(1-\alpha-\beta) &= \omega + E(\underbrace{E(v_t|I_{t-1})}_{=0}) - \beta E(\underbrace{E(v_{t-1}|I_{t-1})}_{=0}) \\
\sigma_t^2 &= \frac{\omega}{(1-\alpha-\beta)}
\end{aligned} \tag{4}$$

In Equation 3, it is stated that $E(v_t|I_{t-1}) = 0$. Using the law of iterated expectations (LIE), we can show that also the unconditional expectation of $E(v_t) = 0$ (as the unconditional expectation is also zero when the conditional expectation is zero). This leads to a unconditional variance, which is only defined ($< \infty$) if the denominator $(1-\alpha-\beta) > 0$, hence if $\alpha + \beta < 1$. This is the requirement for y_t being stationary.

3.2 GARCH Extensions

The formulation of the previously described GARCH model in Equation 2 can be arbitrarily extended. We consider the following extensions:

AGARCH The variance for the asymmetric GARCH (AGARCH[3]) is formulated as:

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \gamma\epsilon_{t-1} + \beta\sigma_{t-1}^2 \qquad \text{or} \qquad \sigma_t^2 = \omega' + \alpha(\epsilon_{t-1} + \gamma')^2 + \beta\sigma_{t-1}^2 \tag{5}$$

where a negative γ implies positive shocks to result in a smaller increase of volatility. The conditional variance must still be positive, therefore requiring $\gamma' > 0$ and $\alpha, \beta \geq 0$.

EGARCH The variance for the exponential GARCH (EGARCH[4]) is formulated as:

$$log(\sigma_t^2) = \omega + \alpha\epsilon_{t-1}^2 + \gamma(|z_{t-1}| - E(|z_{t-1}|)) + \beta log(\sigma_{t-1}^2) \tag{6}$$

This extension explicitly allows for an asymmetry interplay of return and volatility. The volatility equation of the EGARCH uses the standardized residuals of the initial GARCH model (Equation 2) denoted as $z_t = \epsilon_t\sigma_t^{-1}$. For a negative γ, a negative shock will have a bigger impact on future volatility,

[3] Glossary to ARCH (GARCH),Tim Bollerslev,CREATES Research Paper 2008-49.
[4] Glossary to ARCH (GARCH),Tim Bollerslev,CREATES Research Paper 2008-49.

than positive shocks of the same magnitude. Using the log-transformation of the conditional variance, we can avoid any struggles with having to constrain the coefficients of the variance equation to be (strict) positive. The conditional variance $\sigma_t^2 = exp(...)$ will always be positive, regardless of its input into the exp-function.

3.3 Estimator MLE

Stating our GARCH model in a more general manner with $y_t = \mu_t + \epsilon_t$, and $\epsilon_t = \sigma_t z_t$, we can state μ_t and σ_t^2 as function parameters θ. Again we assume $z_t \sim iid.(0,1)$ By assuming Gaussianity for z_t, we can estimate the model parameters θ by using the log-likelihood function and deriving the Maximum-Likelihood-Estimator (MLE): In case z_t is non-Gaussian, which is much more realistic, the log-likelihood-function still gives an consistent estimator. However this one is no longer the most efficient estimator and thus the Quasi-Maximum-Likelihood-Estimtor (QMLE). The $\hat{\theta}_{QMLE}$ satisfies the requirements for consistency (I) and asymptotic normality (II):

$$\hat{\theta}_{QMLE} = arg\ max \sum_{t=1}^{T} log\ell(\theta): \quad (I)\ \hat{\theta}_{QMLE} \xrightarrow{p} \theta_0; \quad (II)\ \sqrt{T}(\hat{\theta}_{QMLE} - \theta_0) \xrightarrow{d} N(0, I^{-1}JI^{-1}) \quad (7)$$

The $\hat{\theta}_{QMLE}$ uses the sandwich formula for the variance with I being the information matrix and J the squared residuals.
To estimate an EGARCH in case of z_t being non-Gaussian, we need a different estimation-function, namely the generalized error distribution (GED), which has more probability mass in the tails. The GED includes an additional parameter v which estimates said fatness of the tails (fatter tails than Gaussian for $0 < v < 2$, GED = Gaussian for $v = 2$ and thinner tails for $v > 2$).

3.4 Misspecification tests

Our final model should no longer have any ARCH-effects. Because the ARCH test has however also power against residual autocorrelation (i.e. autocorrelation in ϵ_t always implies autocorrelation in ϵ_t^2, but not vice-versa), we have to test for no-autocorrelation first.

To test the null-hypothesis of no autocorrelation a Portmanteau test uses $\rho_j = \frac{\sum_{t=j+1}^{T}(\hat{\epsilon}-\bar{\epsilon})(\hat{\epsilon}_{t-j}-\bar{\epsilon})}{\sum_{t=j+1}^{T}(\hat{\epsilon}-\bar{\epsilon})^2}$ to derive the test statistic $\xi_{LB} = T^2 \sum_{j=1}^{h} \frac{\rho_j^2}{T-j}, \underset{d}{\rightarrow} \chi^2(h-p)$. Under the null ξ_{LB} is χ^2 distributed (with p being the sum of AR and MA terms).

As ARCH(p)-test we use a Breusch-Pagan LM-test with the null-hypothesis of No-ARCH effects. Formally speaking, this is H_0: $\gamma_1 = \gamma_2 = ... = \gamma_p = 0$ when running $\hat{\epsilon}_t^2 = \gamma_0 + \hat{\epsilon}_{t-1}^2\gamma_1 + \hat{\epsilon}_{t-2}^2\gamma_2 + ... + \hat{\epsilon}_{t-p}^2\gamma_p + v_t$. The test statistic distributed as $\xi_{ARCH} = T \cdot R^2 \xrightarrow{d} \chi^2(p)$.

Besides, we also test for residual normality by applying a Jarque-Bera test. This test has the null-hypothesis that the error terms are normally distributed and $\xi_{JB} = \xi_S + \xi_K \xrightarrow{d} \chi^2(2)$.

4 Empirical Analysis

When modelling the volatility of returns of an asset one should also consider the trade volume (TV) of the underlying asset at that time. Empiric literature suggests, that there is a positive correlation between the volatility of an asset's returns and the trade volume.[5][6] Wagner and Marsh argue in their paper, that not all trading volume affects the volatility equally, as a great part of trading volume is predictable. In example if an asset was traded intensively yesterday, then it tends to be also traded intensively today too (see our auxiliary regression of trade volume in Equation 8. The authors conclude that is the part of trade volume which is unpredictable - surprise trade volume (SPV) - and significantly enhances the model performance.

4.1 Surprise Trading Volume

To examine the SPV, we conduct an auxiliary regression that explains the daily TV. In line with Wagner and Marsh (2005) we define abnormal logarithmic volume as the difference between the logs of actual and normal volume, using a 50-day moving average for the normal volume.

$$S_t = logVol_t - logVolMA50_t$$

As it can be seen from the ACF plot in Figure 2, this variable is highly autocorrelated. We therefore aim to filter out the predictive part by running the following simple univariate time series model.

$$S_t = \delta + \sum_{i=1}^{k} \phi_i S_{t-i} + \gamma_1 FirstTrDay_t + \gamma_2 FirstTrDay_t + \kappa_1 isJan_t + ... + \kappa_{11} isNov_t + f_t \qquad (8)$$

We start with including one lag of the dependent variable and increase this up to six lags which seems to be a reasonable model that does not suffer from autocorrelation in the residuals. A Breusch-Godfrey LM test with the null hypothesis of No-Autocorrelation is clearly not rejected. The same insight is reflected by the ACF plot, see Figure 2. We store the estimated residuals $\hat{f}_t$ which we from now on call surprise trading volume (SPV). Because we do not make any inference with for the trading volume, it is sufficient for the model to be well specified. Therefore we do not report any detailed results here.

4.2 Volatility of DAX 30 returns

There are two main aspects of analysing the behaviour of the DAX 30 returns: the conditional mean and the conditional variance. The ACF of the returns in Figure 1 shows that there is close to no auto correlation. Hence previous observations give us no predictive power for today's observation. To model the conditional variance with an GARCH-type model, one can choose between a huge amount of different models[7].

First we examine whether or not positive and negative innovations have a different impact on the conditional variance by estimating an AGARCH(1,1). This reveals significant asymmetry. Considering this and the fact that the inclusion of negative SPV could lead to negative conditional variance in case of a positive coefficient, we decide to use an EGARCH.

Because we want to come up with a model that explains all changes in the variance, we include as many lags as needed to avoid any potential ARCH or GARCH effects. We use a general-to-specific modeling approach to identify the model. Starting with an EGARCH(6,6) one can eliminate insignificant ARCH and GARCH terms step by step until an EGARCH(5,4). We never assume a normal distribution of the innovations, as there is way too much probability mass in the tails in every following model.

Although the fifth GARCH lag of an EGARCH(5,4) is insignificant we choose to include it because it is significant when including positive or negative SPV (SPVpos or SPVneg). The EGARCH(5,4) is well specified, as no misspecification (Portmanteau and no-autocorrelation) test can be rejected (see (2) in the regression table).

[5] Wagner, Marsh (2005) Quantitative Finance "Surprise volume and heteroskedasticity in equity market returns"
[6] Brailsford (1996) Accounting Finance "The empirical relationship between trading volume, returns and volatility"
[7] Tim Bollerslev "Glossary to ARCH (GARCH)" 2008

4.3 Volatility of DAX 30 returns with Surprise Trading Volume

It is likely that SPVneg has a different effect on DAX volatility than SPVpos. Wagner and Marsh focus only on SPVpos. As mentioned before, our EGARCH approach has the advantage, that when a variable has a negative effect on variance the variance can only converge to zero for an infinitely high negative effect. The EGARCH conditional variance by construction cannot be negative. Hence, we can extend upon Wagner and Marsh's idea and examine the SPV effects on another level and use the explanatory power of negative surprise trading volume. The regression table contains different specifications of the EGARCH(5,4) and a GARCH(1,1) as a comparison.

The likelihood value increases when including the SPV variables. Especially SPVpos has high explanatory power for both the conditional mean and the conditional variance of the DAX returns. When including SPVpos in the conditional variance the conditional mean can then be modelled as an AR(1) process. In addition to that, it also gives predictive power in the conditional mean. SPVneg itself has not such strong power for explaining both the conditional mean and variance of the DAX returns. The GARCH(1,1) shows that SPVpos has also high explanatory power for a naive model that is not well specified.

The strictly positive conditional variance in the EGARCH is also necessary when including lags of SPVpos, as those lags can have negative coefficients. Including lags of SPV is not necessarily different from including "contemporaneous" SPV's. The surprises that take place in the morning of a trading day, effect the returns in the evening of the trading day. SPV's of prior days significantly affect the conditional variance but not the mean. The estimates for the degrees of freedom of the underlying error distribution increase when including SPV. They increase more when using SPVpos than when using SPVneg (compare (4) and (5)). An increase in the degrees of freedom towards 2 means that the distribution is advancing towards a normal distribution. Specification (7) maximises the likelihood value and minimises the information criteria. The scaled residuals plot of model (7) shows that all volatility clustering is resolved. The GED of (7) is also closest to a normal distribution with 1.89 degrees of freedom.

If there is a positive shock in trading volume (a positive SPV) the DAX returns are significantly more volatile and the expected return of the same day decreases significantly. A negative shock to the traded volume also affects the volatility negatively but does not affect the expected return significantly. SPVpos of past days also don't affect the return's mean significantly. Controlling for shocks also enables us to predict the returns to a certain extend with the return of the previous day with an AR(1) term. Briefly speaking: on average, a high return today leads to a significantly lower return tomorrow.

5 Discussion and Conclusion

Our results are in line with the findings of Wagner and Marsh[8]. We also found that the inclusion of SPV helps explaining both variance and kurtosis of the returns. In addition to the negative impact of positive contemporaneous SPV shocks on the variance we found a negative impact on the conditional mean of the DAX returns as well, raising the model's predictive power on another level. Our EGARCH approach allowed us to use the significant explanatory power of negative SPV and past positive SPV shocks as well.

We found a group of models that give blurry predictive power, as the precision is bound by the daily data. This predictive power should be treated with caution as predicting asset returns is widely considered impossible. Because the data is daily we cannot further examine the dynamics of trading shocks and their effects on the return volatility and expectation. This would be an interesting basis for further research.

[8]Wagner, Marsh (2005) Quantitative Finance "Surprise volume and heteroskedasticity in equity market returns"

6 Appendix

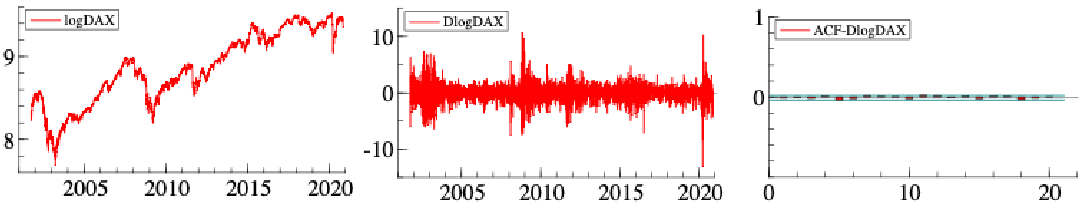

Figure 1: Plot of logDAX, DlogDAX and the ACF of DlogDAX

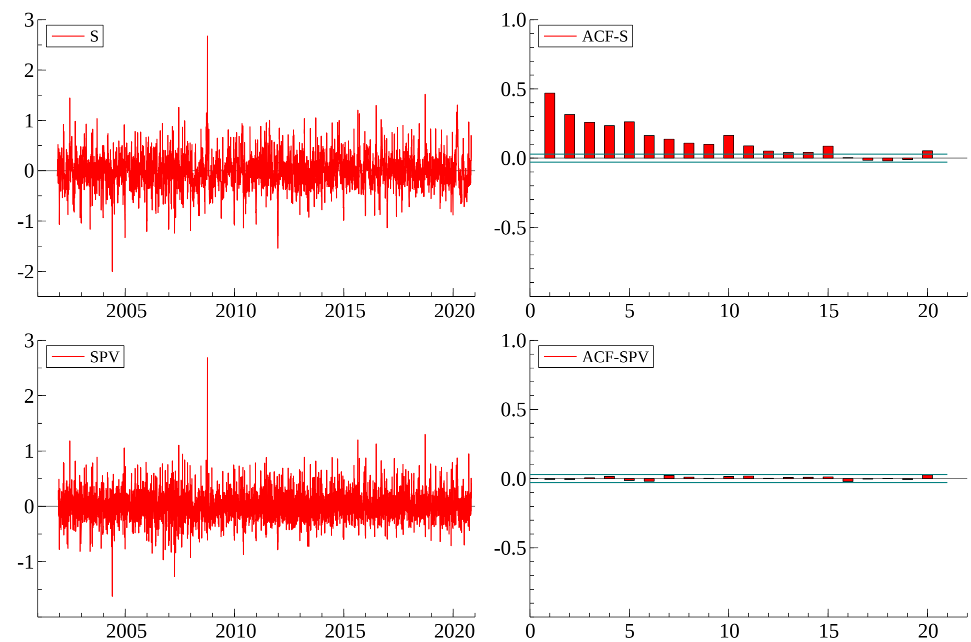

Figure 2: Plots of S and SPV and their ACF

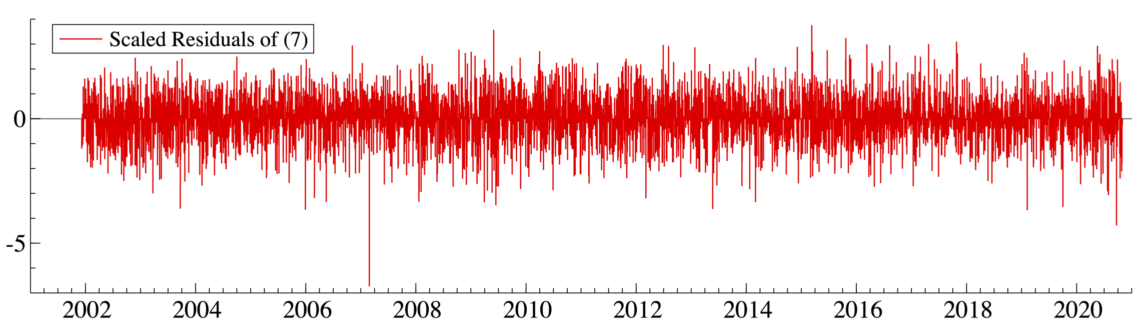

Figure 3: Scaled residuals of model (7)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DlogDAX_1 (Y)	−0.0224	.	−0.0146	−0.0393	−0.0411	−0.0498	−0.0688
	(0.0148)		(0.0125)	(0.0148)	(0.0123)	(0.0117)	(0.0126)
Constant (X)	0.1205	0.0549	0.0570	0.0987	0.0899	0.0971	0.0947
	(0.0159)	(0.0139)	(0.0126)	(0.0156)	(0.0108)	(0.0167)	(0.0133)
SPVpos (X)	−0.5270	.	.	−0.7894	−0.5575	−0.7783	−0.7305
	(0.1110)			(0.1600)	(0.1510)	(0.0612)	(0.1480)
SPVneg (H)	.	.	.	.	2.3620	1.5780	2.0340
					(0.1750)	(0.1800)	(0.1610)
SPVneg_1 (H)	.	.	.	.	.	.	−3.2910
							(0.3370)
SPVneg_2 (H)	.	.	.	.	.	.	2.5590
							(0.4000)
SPVneg_3 (H)	.	.	.	.	.	.	−2.7780
							(0.2920)
SPVneg_4 (H)	.	.	.	.	.	.	1.5000
							(0.1890)
SPVpos (H)	0.6029	.	.	2.3580	.	2.0180	2.5610
	(0.129)			(0.1560)		(0.1510)	(0.1420)
SPVpos_1 (H)	.	.	.	.	.	.	−3.7640
							(0.2600)
SPVpos_2 (H)	.	.	.	.	.	.	2.9870
							(0.2040)
SPVpos_3 (H)	.	.	.	.	.	.	−3.8860
							(0.2430)
SPVpos_4 (H)	.	.	.	.	.	.	2.1550
							(0.1670)
alpha_0 (H)	0.00005	−0.0001	−0.0002	−0.1968	0.2228	−0.0284	−0.0025
	(0.0003)	(0.0004)	(0.0004)	(0.0152)	(0.0194)	(0.0228)	(0.0012)
alpha_1 (H)	0.1102	.	.	.	.	.	.
	(0.0138)						
alpha_2 (H)	.	0.6597	0.6596	1.2720	1.4070	1.2910	−0.7096
		(0.0357)	(0.0378)	(0.3170)	(0.2460)	(0.3190)	(0.1060)
alpha_3 (H)	.	−0.5893	−0.5862	1.0520	1.2010	1.0570	0.6298
		(0.0322)	(0.0338)	(0.2940)	(0.2600)	(0.2900)	(0.1450)
alpha_4 (H)	.	−0.9335	−0.9309	0.5911	0.7112	0.4851	−0.8503
		(0.0396)	(0.0407)	(0.2300)	(0.2310)	(0.2390)	(0.0781)
beta_1 (H)	0.8578	0.3741	0.3561	0.0364	−0.2072	0.0386	1.682
	(0.0182)	(0.1260)	(0.1360)	(0.0691)	(0.0911)	(0.0716)	(0.0766)
beta_2 (H)	.	1.1260	1.128	0.1265	0.0505	0.0721	−1.3710
		(0.0521)	(0.0557)	(0.0603)	(0.0606)	(0.0871)	(0.1110)
beta_3 (H)	.	0.3615	0.3805	0.1929	0.4494	0.3146	1.6180
		(0.1490)	(0.1610)	(0.0877)	(0.0602)	(0.1020)	(0.0819)
beta_4 (H)	.	−0.9327	−0.9227	0.2568	0.4133	0.2386	−1.0230
		(0.0561)	(0.0612)	(0.0632)	(0.0939)	(0.0879)	(0.0894)
beta_5 (H)	.	0.0696	0.0566	0.3270	0.2130	0.2772	0.0924
		(0.1230)	(0.1330)	(0.0493)	(0.0434)	(0.0518)	(0.0452)
eps[-1] (H)	.	−0.1561	−0.1591	−0.0875	−0.1018	−0.0793	−0.1098
		(0.0230)	(0.0244)	(0.0212)	(0.0216)	(0.0199)	(0.0167)
—eps[-1]— (H)	.	0.1305	0.1291	0.1243	0.1410	0.1290	0.1105
		(0.0156)	(0.0154)	(0.0257)	(0.0221)	(0.0261)	(0.0128)
GED log(nu/2)	.	−0.3206	−0.3315	−0.1718	−0.2619	−0.1453	−0.0531
		(0.0324)	(0.0327)	(0.0339)	(0.0386)	(0.0398)	(0.0438)
student-t df	8.6880	.	.	.	.	.	.
	(1.1300)						
alpha(1)+beta(1)	0.968	0.135	0.141	3.855	4.237	3.774	0.069
Log-lik.	-7589.136	-7648.728	-7525.387	-7340.654	-7414.716	-7287.309	-7169.256
AIC	3.167	3.158	3.142	3.066	3.097	3.044	3.001
HQ	3.171	3.164	3.149	3.073	3.104	3.052	3.013
SC/BIC	3.178	3.175	3.161	3.088	3.118	3.067	3.035
Portmanteau, 1-69	[0.78]	[0.81]	[0.86]	[0.64]	[0.65]	[0.58]	[0.68]
No ARCH(1-7)	[0.006]	[0.151]	[0.134]	[0.119]	[0.356]	[0.456]	[0.02]
Normality	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
T	4798	4853	4799	4799	4799	4799	4794
Sample start	2001-12-04	2001-09-18	2001-12-03	2001-12-03	2001-12-03	2001-12-03	2001-12-10
Sample end	2020-10-30	2020-10-30	2020-10-30	2020-10-30	2020-10-30	2020-10-30	2020-10-30

Table 1: The table shows estimates of the different GARCH models applied. Dependant variable is DlogDAX. Y and X refer to regressors of the mean, H refer to regressors of the variance. Standard errors in $(\cdot)$ and p-values in $[\cdot]$ for misspecification tests.