VRIJE UNIVERSITEIT AMSTERDAM

EMPIRICAL FINANCE

Case 3

Authors: **Group 56**
Soheel Gholamy: 2668182
Mark de Kwaasteniet: 2649271

S.gholamy@student.vu.nl
M.G.de.Kwaasteniet@student.vu.nl

Supervisors:
Anne Opschoor
Eva Mynott

Amsterdam, October 18, 2021

1 Question 1a

The data set contains daily open-to-close returns of a particular U.S. stock and the Realized Variance (RV) from 2001 until December 2018 in percentages. The time series contains 4472 observations. The Skewness and Kurtosis test for the variables in Table 1 indicates that the probability that each of them, jointly, is normally distributed, is lower than 1%.

To detect outliers, box-plots of the variable Return and RV were created. Return does not contain significant outliers that need to be treated, since they aren't extreme and therefore could also be a shock in the stock price instead of a data error. However, the variable RV contains outliers which we treat by winsorizing on the 0.1% level from the upper side. The summary statistics of the original variables are in Table A that can be found in the Appendix. The adjusted summary statistics are presented in Table 1.

Table 1

Summary Statistics

Variable Names	Mean	Std. Deviation	Min	Max	P-Value
Return	-0.003	1.404	-9.0844	8.8101	0.000
Realized Variance (RV)	2.436	3.462	0.113	45.843	0.000

This table contains the adjusted summary statistics of the variables in the data set of case 2 - group 56. The Realized Variance (RV) is winsorized on the 0.1% level from the upper side. The mean, standard deviation, minimum value and maximum values are represented in column 2,3,4 and 5 respectively (N = 4472, which are daily observations). P-value denotes the probability that the Chi-Squared distributed variables is greater than the test statistic. The p-value is calculated with the Jarque Bera test, which jointly tests the skewness and kurtosis. It represents the probability of accepting the null-hypothesis of the test, which is a joint hypothesis of the skewness being equal to zero and kurtosis being equal to three.

The annualized return volatility can be calculated from the standard deviation of Return. There are 4472 observations in 18 years, meaning that, on average, there are 248.44 trading days in a year. Therefore, we assume 248 trading days for our calculations.

$$Annualized - Volatility = 1.404\% \times \sqrt{248} = 22.11\% \tag{1}$$

The annualized return volatility can be calculated from the mean of the Realized Variance. We assume 248 trading days.

$$Annualized - Volatility = \sqrt{2.346\%} \times \sqrt{248} = 24.12\% \tag{2}$$

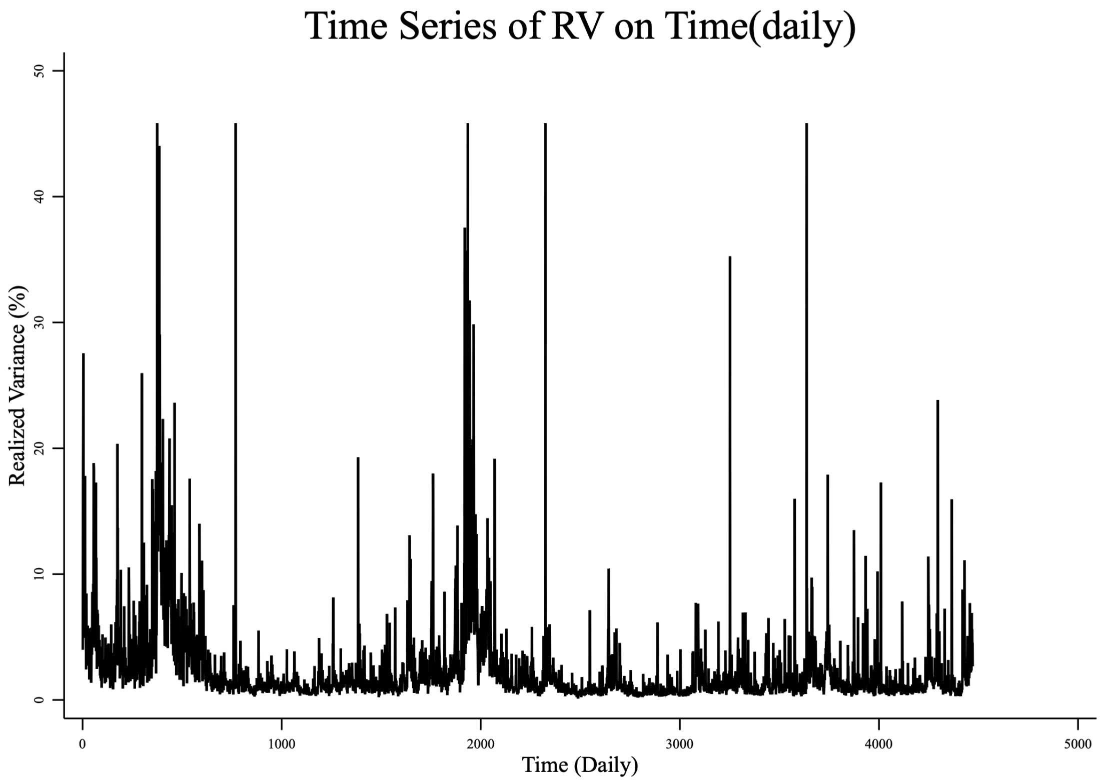

Figure 1: This figure contains the time series plot of the Realized Variance (RV) at day t in percentages. Time (t) is the daily measure from 2 January 2001 to 31 December 2018.

2 Question 1b

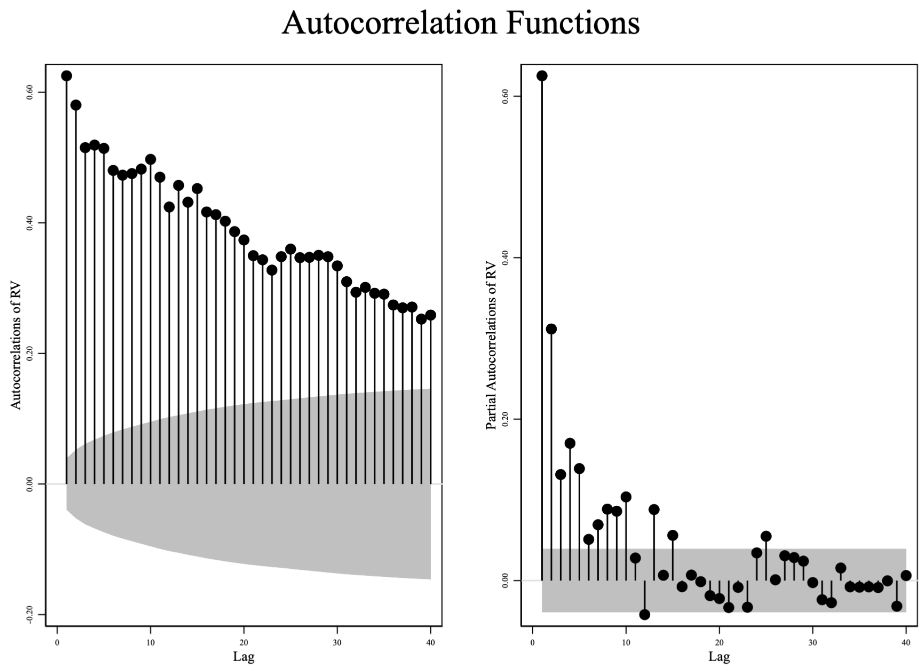

Figure 2: Left: the autocorrelations of RV (Realized variance at day t). Right: The partial autocorrelations for RV (Realized variance at day t). Time ($daily$) is from 2 January 2001 up to 31 December 2010. The 95% confidence bands are calculated using Barlett's formula for MA(q), i.e., SE = $\frac{1}{\sqrt{n}}$

.

To indicate a first guess of the ARMA(p,q) model, we estimate the orders p and q by using the (partial) autocorrelation functions. AR(p) models have theoretical PAC functions with non-zero values at the AR terms in the model and zero values otherwise. The AC function will exponentially decline to zero. We observe that in our time series, there is significant high partial autocorrelation in the first lag. Moreover, we observe that the autocorrelation plot is not exponentially declining to zero, thus this makes it difficult to determine what lag to use. However, since we don't want to overfit the model, we use a lag of 1 for the Autoregressive Process.
MA(q) models have a theoretical ACF with non-zero values at the MA terms in the model and zero values elsewhere. The PACF curve will exponentially decline to zero. We observe from the AC and PACF curve that there is a slow decline over the lags. Therefore, we conclude the need of an ARMA(p,q) model. We use 3 lags for the Moving Average Process, because there is a fast decline visible in the partial autocorrelation after the second and third lag.

Conclusion: By looking at the ACF and PACF, we guess that using an ARMA(1,3) might be suitable for our RV time series.

3 Question 1c

We estimate two ARMA(p,q) models for RV using the in-sample observations between 2001 and 2010:

- **RV Model I:** ARMA(3,1) model
- **RV Model II:** ARMA(1,3) model

RV Model II, ARMA(1,3):

$$RV_t = c + \phi_1 RV_{t-1} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \theta_3 \epsilon_{t-3} + \epsilon_t \tag{3}$$

Table 2

Estimation of both ARMA RV models

VARIABLES	RV Model I	RV Model II
AR		
ϕ_1	1.161***	0.985***
ϕ_2	-.095	
ϕ_3	-.077	
MA		
θ_1	-0.861***	-0.677***
θ_2		-0.036***
θ_3		0.098***
Constant RV	3.156**	3.156**
Constant σ	2.937***	2.934***
Observations	2,490	2,490
AIC	12445.03	12439.23
BIC	12479.95	12474.20
Maximum Likelihood	-6216.51	-6213.64

Table 2 contains the ARMA(p,q) models for RV using the in-sample from 2 January 2001 up to 31 December 2010. The second column denotes the calculated estimators and statistics AIC, BIC and Maximum likelihood) for Model 1 The second column denotes the calculated estimators and statistics for Model 2. Statistical significance at the 1%, 5% and 10% levels are denoted by ***, ** and *, respectively (N = 2490).

We evaluate RV Model I and RV Model II by comparing the AIC, BIC and Maximum Likelihood of the two models.

$$\textbf{AIC(k)} = -2\log\left(\hat{L}\right) + 2k \tag{4}$$

$$\textbf{BIC(k)} = -2\log\left(\hat{L}\right) + k\log T \tag{5}$$

where T, k and $\hat{L}$ denote the sample size, number of parameters and the maximum likelihood, respectively.

RV Model II has the lowest AIC and BIC. Moreover, its Maximum Likelihood is higher than that of RV Model I. This is the likelihood that the model produces the same outcomes as the data that is actually observed, meaning that a higher number is desirable. This suggests that RV Model II nicely straddles the requirements of goodness-of-fit and parsimony. We can conclude that the ARMA(1,3) model offers a better fit than ARMA(3,1), while at the same time not losing parsimony imposed by including the additional MA lags. In conclusion, our favorite model is RV Model II – ARMA(1,3)

4 Question 1d

To test if the ARMA models are misspecified, we perform the multivariate Ljung-Box (Q) test for white noise. This test is applied to the residuals of RV Model I and II. It tests if there is significant autocorrelation for lags one through lag k. To determine the lag length, we use the AC function on the residuals of the ARMA models. These AC functions, figure 5 in the appendix, show us that we should use 14 lags in the Ljung-Box test formula ($k = 14$).

$$\mathbf{LB}(m) = T(T+2)\sum_{t=1}^{m} \frac{r_k(\hat{\epsilon})^2}{T-k} \sim \chi^2(m) \tag{6}$$

where

$$r_k(\hat{\epsilon}) = \frac{\sum_{t=k+1}^{T} \hat{\epsilon}_t \hat{\epsilon}_{t-k}}{\sum_{t=1}^{T} \hat{\epsilon}_t^2} \tag{7}$$

where T, k and m denote the sample size, number of parameters and the degrees of freedom, respectively.

Hypothesis Set 1

- H_0: There is no significant autocorrelation up to lag 14.
- H_A: There is significant autocorrelation up to lag 14.

Results

- **ARMA(3,1) Model:** $Q_{statistic} = 49.01$
- **ARMA(1,3) Model:** $Q_{statistic} = 43.40$
- **Distribution** $\chi^2(5)$**:** $Q_{critical} = 11.07$

Model I and Model II have both a higher Q_{stat} than the Q_{crit} at 5% significance level. Therefore, we reject the null hypotheses, meaning that there is significant autocorrelation left in the residuals of Model I and Model II up to the 14th lag. In conclusion, the statistics indicate that both the models are misspecified.

We notice, by using various lags in the LB-test, that we cannot reject the null-hypothesis up till lag 9. Meaning that there is no significant autocorrelation up to lag 9. However, if we increase the number of lags beyond 9, then the Q_{stat} becomes larger than the Q_{crit}. This is inline with our AC function in the appendix.

5 Question 1e

Hypothesis Set 2

- H_0: Realized Variance RV contains a unit root.
- H_A: Realized Variance RV has a stationary process.

We analyze the plot of the Realized Variance on Time (Figure 1) to observe if we should add a constant to the model. Indeed, it looks like the Realized Variance has an unconditional mean since it always declines to a specific value after a peak. Therefore, a constant is added. We do not add a trend term, since it is observable in Figure 1 that the Realized Variance does not increase or decline over time. There is no drift visible.

The Augmented Dickey Fuller (ADF) test with 14 lags (p = 14):

$$\Delta RV_t = \alpha + \psi_1 RV_{t-1} + \psi_2 \Delta RV_{t-1} + \psi_3 \Delta RV_{t-2} + \cdots + \psi_{15} \Delta RV_{t-14} + \epsilon_t \tag{8}$$

- H_0: $\psi_1 = 0$
- H_A: $\psi_1 < 0$

$$\text{DF}_\tau = -6.892 \tag{9}$$

$$\text{DF}_{crit,5\%} = -2.860 \tag{10}$$

Based on the $DF_{statistic}$ and the $DF_{critical}$, we can reject the null hypothesis. This means that there is no unit root found in the Realized Variance time series and that it is indeed a stationary process.

This conclusion makes economically sense since stock returns tend to fluctuate around a specific value. Shocks in the return have a transitory effect which means that the effect is more visible in short-term returns than in long-term returns. Eventually, the effect of the shock on future values dies out. Therefore, the shock is not permanent and the returns have an unconditional mean and variance.

We are analyzing the Realized Variance, which is essentially the variance in the stock returns. Therefore, the shocks in Realized Variance have a transitory effect and we expect to see a stationary process.

6 Question 1f

Table 3

Estimation of GARCH and E-GARCH

VARIABLES	Model 3 - GARCH	Model 4 - E-GARCH
L.earch		-0.0175**
		(0.00883)
L.earch_a		0.134***
		(0.0126)
L.egarch		0.993***
		(0.00227)
L.arch	0.0571***	
	(0.00638)	
L.garch	0.939***	
	(0.00638)	
Constant Return	-0.0132	-0.0301
	(0.0230)	(0.0234)
Constant Sigma	0.00904**	0.00693***
	(0.00361)	(0.00212)
Observations	2,490	2,490
AIC	8578	8582.64
BIC	8601.28	8611.74
Maximum likelihood	-4285	-4286

Table 3 contains the GARCH(1,1) and EGARCH(1,1) models for Return using the in-sample from 2 January 2001 up to 31 December 2010. The second column denotes the calculated estimators and statistics (AIC, BIC and Maximum likelihood) for Model 3. The second column denotes the calculated estimators and statistics for Model 4. Statistical significance at the 1%, 5% and 10% levels are denoted by ***, ** and *, respectively. (N = 2490).

Model equation of E-GARCH(1,1):

$$Return_t = \mu_t + \epsilon_t,\ \epsilon \sim N(0, \sigma_t^2) \tag{11}$$

$$ln(\sigma_{t+1}^2) = \omega + \alpha\left(\frac{|\epsilon_t|}{\sigma} - \sqrt{\frac{2}{\pi}}\right) + \gamma\frac{\epsilon_t}{\sigma_t} + \beta \ln(\sigma_t^2) \tag{12}$$

We evaluate Model 3 and Model 4 by comparing the AIC, BIC and Maximum Likelihood of the two models. The GARCH model has the lowest AIC, BIC and highest Maximum Likelihood. Therefore, we observe that the assumption for the EGARCH(1,1) model that negative shocks at time t-1 have a stronger impact in the variance at time t than positive shocks, does not offer a model with improved fit. In conclusion, our preferred model is the GARCH(1,1).

7 Question 1g

α**:** is the effect of the scaled absolute news deviated from its mean on the natural log of the conditional variance σ^2_{t+1}. The α is used to describe the possible volatility clusters in the data, meaning that it should be positive. In that case if the absolute value of ϵ_t is large we possibly have a high volatility cluster and thus we want the next conditional variance σ^2_{t+1} to be large as well. The estimation results of our EGARCH model denote that the coefficient of α is indeed positive (α = 0.134), thus this result is in line with our expectations.

β**:** is the effect of the σ^2_t on the natural log of the conditional variance σ^2_{t+1}. Again we expect the β to be positive ($\beta \geq 0$), since volatility should be positively related to each other for autocorrelation. Furthermore, we expect a high coefficient for the β, since we the expect the variance to be highly autocorrelated. The estimation results of our EGARCH model denote that the coefficient of β is indeed positive and very close to 1 (β = 0.993), supporting the clusters of volatility which is in line with our expectations.

γ**:** is the extra effect of negative news on the natural log of the conditional variance σ^2_{t+1}. Meaning that if ϵ_t is negative, the value of the volatility should increase more compared to if ϵ_t is positive. Therefore, we expect the γ to be negative. In this case negative times negative becomes positive, and it will have a stimulating effect on the volatility. The estimation results of our EGARCH model denote that the coefficient of γ is indeed negative (γ = -0.0175), thus this result is in line with our expectations.

We cannot compare the log likelihood/AIC/BIC values of Table 3 with those of Table 2. Table 3 contains the GARCH and EGARCH models that are estimated using the returns and Table 2 contains the ARMA models that estimated using the Realized Variance. Therefore, we cannot compare the aforementioned statistics of these tables, because the dependent variable in the GARCH and EGARCH models differ from the dependent variable in the ARMA models. If the dependent variables in the GARCH and ARMA models were both Returns or Realized Variances, then it would have been possible to compare.

8 Question 1h

Derivation of the one-step forecast of the ARMA(1,3) model:

$$\begin{aligned} RV_{T+1} &= \mathbb{E}(RV_{T+1}|\Omega_T) && (13)\\ &= \mathbb{E}(c + \phi_1 RV_T + \epsilon_{T+1} + \theta_1\epsilon_T + \theta_2\epsilon_{T-1} + \theta_3\epsilon_{T-2}|\Omega_T) && (14)\\ &= c + \phi_1 RV_T + \theta_1\epsilon_T + \theta_2\epsilon_{T-1} + \theta_3\epsilon_{T-2} && (15) \end{aligned}$$

Derivation of the one-step forecast of the GARCH(1,1) model:

$$\begin{aligned} \sigma^2_{T+1} &= \mathbb{E}(\sigma^2_{T+1}|\Omega_T) && (16)\\ &= \mathbb{E}(\omega + \alpha\epsilon^2_T + \beta\sigma^2_T|\Omega_T) && (17)\\ &= \omega + \alpha\epsilon^2_T + \beta\sigma^2_T && (18)\\ &= \sigma^2_{T+1} && (19) \end{aligned}$$

$$Return_T = \mu + \epsilon_T = \mu + \sigma_T u_T,\ u_T \sim N(0,1) \tag{20}$$

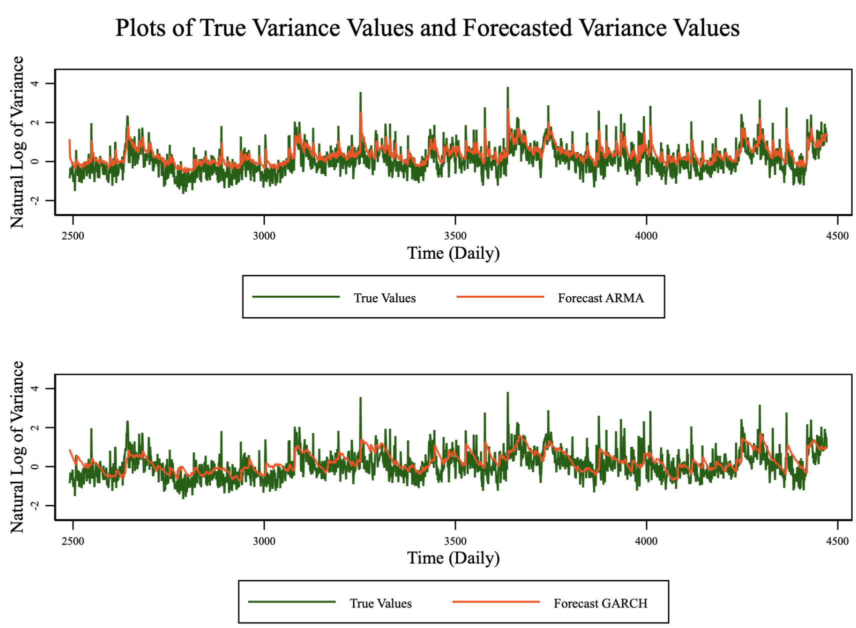

Figure 3: Twoway line plot including the natural log of the True Value and the natural log of the forecasted values of the ARMA(1,3) and GARCH(1,1) model.

If we look at both the natural log forecasted values of the ARMA(1,3) and GARCH(1,1) model, then the only remarkable difference between the two forecasted time series is that the ARMA model has, noticeably, higher peaks compared to the GARCH model. This is probably due to the fact that the Realized Variance variable contained more extreme observations compared to the Returns variable.

9 Question 1i

Testing on unbiasedness

To test whether the 1-step ahead forecasts of the two models are unbiased, we generate the forecast error by subtracting the forecasted values of the variance from the true values (Realized Variance) in the out of sample period.

To determine whether the average forecast error is significantly different from zero, we regress the the generated forecast error on a constant. However, before we can regress the forecast error, we need to know whether there is autocorrelation present in the forecast error. For both models this holds true, using an AC function to determine this. Therefore, we use Newey West Standard errors in the regression of the forecast error on a constant.

Autocorrelation of forecast errors estimated with ARMA(1,3): 2 lags

Autocorrelation of forecast errors estimated with GARCH(1,1): 3 lags

We test hypothesis set 3, determining the unbiasedness of the forecast errors, using the t_{stat} in equation 21. We use the t_{crit} of a 5% significance level to determine whether we should accept or reject the null-hypothesis. Table 9 contains the results of the unbiasedness t-test of both the ARMA(1,3) and the GARCH(1,1) model. We conclude that the forecast errors of the ARMA(1,3) are biased, rejecting the null-hypothesis, meaning that the average forecast error is significantly different from zero. However, the forecast errors of the GARCH(1,1) model are unbiased, accepting the null-hypothesis, meaning that the average forecast error is zero.

Hypothesis Set 3

- H_0: $e_{t+1|t} = 0$
- H_A: $e_{t+1|t} \neq 0$

$$t_{stat} = \frac{\hat{e}_{t+1|t} - e_{t+1|t}}{S.E.} \sim t(0, 1, N-k) \tag{21}$$

$$t_{crit,5\%} = -1.96 \tag{22}$$

Testing on efficiency

To test on efficiency, we regress the true variance values on the forecasted variance values, forecasted with the ARMA(1,3) and GARCH(1,1) model. Equation 23 denotes the general regression model that is used to perform the efficiency test. To test whether the forecast is efficient, we jointly test the significance of the constant equal to zero and the coefficient for the forecasted variance values equal to one. Using the F-test and the F_{crit}, we will test the null-hypothesis in hypothesis set 4. From table 9, we conclude that the we can reject the null-hypothesis in hypothesis set 4, meaning that both the ARMA(1,3) and GARCH(1,1) model don't create an efficient forecast.

The same conclusion arises from looking at the scatter plots in figure 4, where we can see that the is no pattern or relationship noticeable in the data. If the forecasts were efficient, we would expect the forecasted variance values to have a linear relationship with the true variance values. This would look like a linear line from the left lower corner to the right upper corner, however, we clearly see that this is not the case.

$$Realized\_Values(winsorized) = \alpha + \beta_1 Fcst\_Variance\_Values_t + \epsilon_t \tag{23}$$

Hypothesis Set 4

- H_0: $\alpha = 0$ & $\beta_1 = 1$
- H_A: At least one of the estimated parameters is not equal to its predetermined value.

$$F_{crit} \approx 2.70 \sim F_{5\%(2,1980)} \tag{24}$$

Table 4

Unbiasedness and Efficiency Tests

	ARMA(1,3) Model	GARCH(1,1) Model
Unbiasedness		
T-statistic	-3.25	-0.21
P-value	0.001	0.833
Efficiency		
F-statistic	29.74	21.27
P-value	0.000	0.000

This table contains the test statistics and p-values of the performed Unbiasedness and Efficiency tests for the Forecast Errors of the ARMA(1,3) and GARCH(1,1) models.

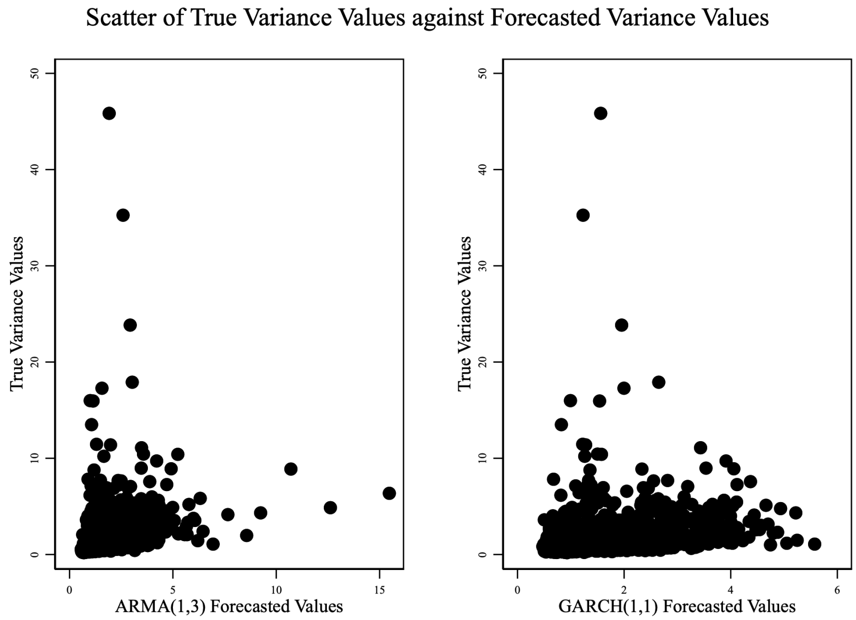

Figure 4: Scatter plots including the True Values against the forecasted values of the ARMA(1,3) and GARCH(1,1) model.

10 Question 1j

Using the statistics we've produced so far, we disagree with Mr. Financial Times Series, that using Realized Measures (which are based on HF-data) leads to better forecasts of the risk of a portfolio of assets than the traditional risk models based on only daily (open-to-close) returns. We temporarily conclude from our forecast evaluations, that the traditional risk models based on only daily returns is superior to using the Realized Variance of the HF-data. This conclusion is primarily based on the fact that the forecasted errors of the GARCH model are unbiased where the forecasted errors of the ARMA model are biased.

11 Question 1k

To test on equal predictive accuracy, we use the Diebold-Mariano test. In order to perform the Diebold-Mariano test, we calculate the difference between the squared forecast errors of the GARCH(1,1) model and the squared forecast errors of the ARMA(1,3) model (D_t). Furthermore, we regress this difference on a constant, using equation 25, to test whether the constant in the regression (average difference), is statistically different from zero. The hypotheses are further specified in hypothesis set 5.
However, before we can regress the forecast error, we need to know whether there is autocorrelation present in the generated difference between the squared forecast errors of the ARMA(1,3) and GARCH(1,1) model. Using the autocorrelation function, we conclude that there is autocorrelation up till the second lag. Therefore, we use

Newey Standard Errors to estimate the coefficient with correct standard errors to calculate the test statistic.

Using the t_{stat} in equation 21 we are able to test the null-hypothesis of the Diebol-Mariano test. We use the t_{crit} of a 5% significance level to determine whether we should accept or reject the null-hypothesis. Equation 26 contains the results of the t-test, concluding that we can reject the null-hypothesis, since the $t_{statistic}$ is larger than the t_{crit}. This means that the models significantly differ from each other in terms of predictive accuracy. We conclude that the ARMA(1,3) model has a higher predictive accuracy compared to the GARCH(1,1) model. Furthermore, comparing the Mean Squared Predictive Errors, we see that the ARMA(1,3) has a lower MSPE, indicating a higher quality of the predicted values.

MSPE Values

- MSPE: GARCH(1,1) Model = 3.848
- MSPE: ARMA(1,3) Model = 3.669

Hypothesis Set 5

- H_0: $\mathbb{E}\left[L(\epsilon_t^{m1})\right] = \mathbb{E}\left[L(\epsilon_t^{m2})\right]$
- H_A: $\mathbb{E}\left[L(\epsilon_t^{m1})\right] \neq \mathbb{E}\left[L(\epsilon_t^{m2})\right]$

$$D_t = \alpha + e_t \tag{25}$$

$$t_{stat} = \frac{0.179}{0.080} \approx 2.25 \tag{26}$$

After testing the predictive accuracy of the ARMA(1,3) and GARCH(1,1) together, we have come to the change our temporary conclusion in question 1j. The ARMA(1,3) has significantly better predictive accuracy compared to the GARCH(1,1) model. Even though, the forecast errors of the GARCH(1,1) are unbiased, we conclude that the ARMA(1,3) model is superior to the GARCH(1,1) model. This conclusion is in line with the plots in figure 3. Using eye-ball econometrics, we notice that the forecasted values of the ARMA model are more similar to the True Variance Values, since the ARMA model creates higher peaks and troughs.

Case Conclusion: We agree with Mr. Financial Times Series, that using Realized Measures (which are based on HF-data) leads to better forecasts of the risk of a portfolio of assets than the traditional risk models based on only daily (open-to-close) returns.

12 Appendix

Table 1

Summary Statistics

Variable Names	Mean	Std. Deviation	Min	Max	P-Value
Return	-0.00336	1.404	-9.0844	8.8101	0.000
Realized Variance (RV)	2.406	5.146	0.113	239.3	0.000

This table contains the summary statistics of the variables in the data set of case 3 - group 56 without treating the outliers. The mean, standard deviation, minimum value and maximum values are represented in column 2,3,4 and 5 respectively (N = 4472). P-value denotes the probability that the Chi-Squared distributed variables is greater than the test statistic. The p-value is calculated with the Jarque Bera test, which jointly tests the skewness and kurtosis. It represents the probability of accepting the null-hypothesis of the test, which is a joint hypothesis of the skewness being equal to zero and kurtosis being equal to three.

The annualized return volatility can be calculated from the mean of the Realized Variance. We assume 248 trading days.

$$Annualized - Volatility = \sqrt{2.406\%} \times \sqrt{248} = 24.43\% \tag{27}$$

Autocorrelation Functions

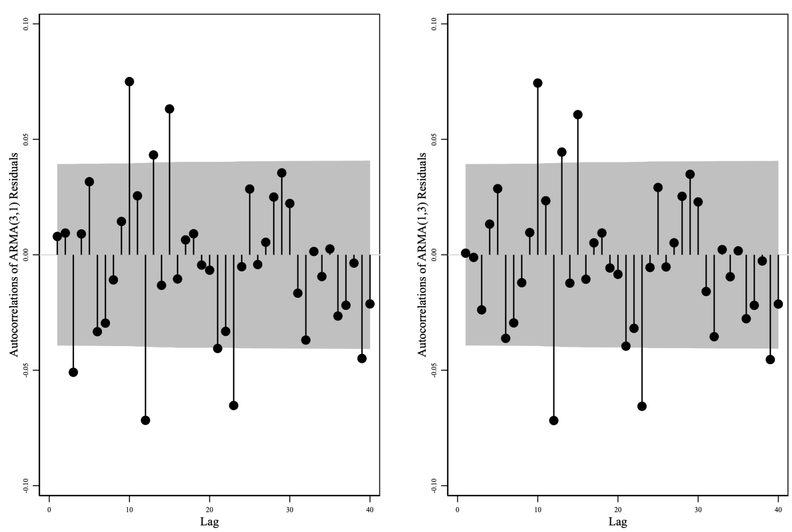

Figure 5: Left: the autocorrelations of the residuals of the ARMA(3,1). Right: the autocorrelations of the residuals of the ARMA(1,3). The 95% confidence bands are calculated using Barlett's formula for MA(q), i.e., SE = $\frac{1}{\sqrt{n}}$

```
//Prepare Workspace
clear all
ssc install outreg2
ssc install asdoc
ssc install winsor

//Change directory
cd "/Users/markdekwaasteniet/Documents/Master Finance/Empirical Finance/Case 3"

//Data structurizing
import excel "/Users/markdekwaasteniet/Documents/Master Finance/Empirical Finance/Case 3/data_case_III_group_56.xls", sheet("Data") firstrow

//Exercise 1a
tsset Time

//Test the normality of all variables except Date
sktest Return RV

mat sktest = r(Utest)

//Subtract the last column of the matrix including the p-values
matrix Pvalues = sktest[.,4...]

//Summarize all the variables except Date
//Sut the variables into a table and convert to matrix
tabstat Return RV, statistic(mean sd min max) columns(statistics) save
mat statistics = r(StatTotal)'

//Merge the two matrices together
mat table1 = statistics,Pvalues\Pvalues,statistics
mat table1 = table1[1..2,1...]

//Output the matrix
asdoc wmat, mat(table1) title(Summary Statistics Table 1) dec(3) replace
//Rename Myfile.doc name to Table 1.doc

//Interpret the data in terms of outliers
summarize Return, detail
graph hbox Return
histogram Return, frequency

summarize RV, detail
graph hbox RV
histogram RV, frequency

//winsorizing the outliers with respect to the 0.1% of the data
winsor RV, p(.001) highonly gen(RV_w1)
graph hbox RV_w1
```

```
//Test the normality of all variables except Date
sktest Return RV_w1

mat sktest = r(Utest)

//Subtract the last column of the matrix including the p-values
matrix Pvalues = sktest[.,4...]

//Summarize all the variables except Date
//Sut the variables into a table and convert to matrix
tabstat Return RV_w1, statistic(mean sd min max) columns(statistics) save
mat statistics = r(StatTotal)'

//Merge the two matrices together
mat table1 = statistics,Pvalues\Pvalues,statistics
mat table1 = table1[1..2,1...]

//Output the matrix
asdoc wmat, mat(table1) title(Summary Statistics Table 1) dec(3) replace
//Rename Myfile.doc name to Table 1 - Adjusted.doc

//Exercise 1b
//Create an in-sample period (including or excluding 2010?)
gen dummy_insample = 0
replace dummy_insample = 1 if Time <= 2490

graph set window fontface "Times New Roman"
line RV_w1 Time, scheme(s2mono) graphregion(color(white)) bgcolor(white)
ytitle("Realized Variance (%)", size(small)) xtitle("Time (Daily)", size(small))
legend(size(vsmall)) yla(, labsize(*0.5) nogrid) xla(, labsize(*0.5) nogrid) title("Time Series
of RV on Time(daily)")

//create the partial autocorrelation function
ac RV_w1 if dummy_insample == 1, scheme(s1mono) graphregion(color(white))
bgcolor(white) ytitle("Autocorrelations of RV", size(small)) xtitle("Lag", size(small))
legend(size(vsmall)) yla(, labsize(*0.5) nogrid) xla(, labsize(*0.5) nogrid) name(ac_RV)
pac RV_w1 if dummy_insample == 1, scheme(s1mono) graphregion(color(white))
bgcolor(white) ytitle("Partial Autocorrelations of RV", size(small)) xtitle("Lag", size(small))
legend(size(vsmall)) yla(, labsize(*0.5) nogrid) xla(, labsize(*0.5) nogrid) name(pac_RV)
graph set window fontface "Times New Roman"
graph combine ac_RV pac_RV, scheme(s1mono) title("Autocorrelation Functions")
graphregion(color(white))

//This looks like an ARMA model
// Using a lag of 2 or 3 because the PACF denotes an fast decline in the autocorrelation after
2 or 3 lags.


//Exercise 1c
arima RV_w1 if dummy_insample == 1, ar(1/3) ma(1)
```

```
outreg2 using "Table 2.tex", replace word addstat(ll, e(ll)) ctitle(RV Model 1 - ARMA(3.1))
title(ARMA estimation results)
estimates store arma31

arima RV_w1 if dummy_insample == 1, ar(1) ma(1/3)
outreg2 using "Table 2.tex", append word addstat(ll, e(ll)) ctitle(RV Model 2 - ARMA(1,3))
title(ARMA estimation results)
estimates store arma13


estimates table arma31 arma13, stats(aic, bic, ll)


// ARMA(1,3) is the best model


//Exercise 1d
arima RV_w1 if dummy_insample == 1, ar(1/3) ma(1)
predict res_RV_arma31 if dummy_insample == 1, residuals
ac res_RV_arma31, scheme(s1mono) graphregion(color(white)) bgcolor(white)
ytitle("Autocorrelations of ARMA(3,1) Residuals", size(small)) xtitle("Lag", size(small))
legend(size(vsmall)) yla(, labsize(*0.5) nogrid) xla(, labsize(*0.5) nogrid)
name(ac_res_ARMA31)
wntestq res_RV_arma31, lags(14)

arima RV_w1 if dummy_insample == 1, ar(1) ma(1/3)
predict res_RV_arma13 if dummy_insample == 1, residuals
ac res_RV_arma13, scheme(s1mono) graphregion(color(white)) bgcolor(white)
ytitle("Autocorrelations of ARMA(1,3) Residuals", size(small)) xtitle("Lag", size(small))
legend(size(vsmall)) yla(, labsize(*0.5) nogrid) xla(, labsize(*0.5) nogrid)
name(ac_res_ARMA13)
wntestq res_RV_arma13, lags(14)

graph set window fontface "Times New Roman"
graph combine ac_res_ARMA31 ac_res_ARMA13, scheme(s1mono) title("Autocorrelation
Functions") graphregion(color(white))

//Exercise 1e
varsoc RV_w1, maxlag(20)
dfuller RV_w1, regress lags(14)

//Exercise 1f
// gen Return_w1_2 = Return_w1^2
// ac Return_w1_2
// pac Return_w1_2

// reg Return_w1_2 L(1/8).Return_w1_2

arch Return if dummy_insample == 1, arch(1/1) garch(1/1)
estimates store GARCH11
```

```
outreg2 using "Table 3.tex", replace word addstat(ll, e(ll)) ctitle(Model 3 - GARCH)
title(GARCH estimation results)
estimates table GARCH11, stats(aic, bic, ll)

// reg Return_w1_2 L(1/8).Return_w1_2
arch Return if dummy_insample == 1, earch(1/1) egarch(1/1)
estimates store EGARCH11
outreg2 using "Table 3.tex", append word addstat(ll, e(ll)) ctitle(Model 4 - EGARCH)
title(GARCH and EGARCH estimation results)
estimates table EGARCH11, stats(aic, bic, ll)


//exercise 1g
//The gamma coefficient is in line with our expectations

//the aic and bic are not comparable, because the dependent variable in the models differ.

//Exercise 1h
arima RV_w1 if dummy_insample == 1, ar(1) ma(1/3)
predict sigma2arma13 if dummy_insample == 0, xb

arch Return if dummy_insample == 1, arch(1) garch(1)
predict sigma2GARCH if dummy_insample == 0, variance

gen ln_RV=ln(RV_w1)
gen ln_fcst_GARCH = ln(sigma2GARCH)
gen ln_fcst_arma = ln(sigma2arma13)

twoway (line ln_RV Time if dummy_insample == 0) (line ln_fcst_arma Time if
dummy_insample == 0),  scheme(s1color) legend(label(1 "True Values") label(2 "Forecast
ARMA")) ytitle("Natural Log of Variance", size(small)) xtitle("Time (Daily)", size(small))
legend(size(vsmall)) yla(, labsize(*0.5)) xla(, labsize(*0.5))
graph save "Forecast ARMA.gph",  replace
twoway (line ln_RV Time if dummy_insample == 0) (line ln_fcst_GARCH Time if
dummy_insample == 0), scheme(s1color) legend(label(1 "True Values") label(2 "Forecast
GARCH")) ytitle("Natural Log of Variance", size(small)) xtitle("Time (Daily)", size(small))
legend(size(vsmall)) yla(, labsize(*0.5)) xla(, labsize(*0.5))
graph save "Forecast GARCH.gph", replace

graph combine "Forecast ARMA.gph" "Forecast GARCH.gph", scheme(s1color) title("Plots
of True Variance Values and Forecasted Variance Values", size(medium))  col(1) iscale(1)
graph save "Exercise 1h.png", replace

twoway (line ln_fcst_arma Time if dummy_insample == 0) (line ln_fcst_GARCH Time if
dummy_insample == 0), scheme(s1color) legend(label(1 "Forecast ARMA") label(2
"Forecast GARCH"))


//Exercise 1i
//Create the forecast error -unbiasedness
```

```
gen ehat_ARMA = RV_w1 - sigma2arma13
gen ehat_ARMA_sq = ehat_ARMA^2
reg ehat_ARMA
ac ehat_ARMA
newey ehat_ARMA, lag(2)
//biased

gen ehat_GARCH = RV_w1 - sigma2GARCH
gen ehat_GARCH_sq = ehat_GARCH^2
reg ehat_GARCH
ac ehat_GARCH
newey ehat_GARCH, lag(3)
//Unbiased

//efficiency test
reg RV_w1 sigma2arma13
test (sigma2arma13=1)(_cons=0)
scatter RV_w1 sigma2arma13 if dummy_insample == 0, scheme(s1color) ytitle("True
Variance Values", size(small)) xtitle("ARMA(1,3) Forecasted Values", size(small))
legend(size(vsmall)) yla(, labsize(*0.5)) xla(, labsize(*0.5))
graph save "forecast scatter arma.gph" , replace
//inefficient

reg RV_w1 sigma2GARCH
test (sigma2GARCH=1)(_cons=0)
scatter RV_w1 sigma2GARCH if dummy_insample == 0, scheme(s1color) ytitle("True
Variance Values", size(small)) xtitle("GARCH(1,1) Forecasted Values", size(small))
legend(size(vsmall)) yla(, labsize(*0.5)) xla(, labsize(*0.5))
graph save "forecast scatter garch.gph", replace
//inefficient
graph combine "forecast scatter arma.gph" "forecast scatter garch.gph", scheme(s1color)
title("Scatter of True Variance Values against Forecasted Variance Values", size(medium))
iscale(1)
graph save "Exercise 1i.png", replace


//exercise 1j
//We disagree with the CEO. The traditional method in our opinion is still the best method to
forecast the volatility of the assets.

//exercise 1k
summarize ehat_GARCH_sq ehat_ARMA_sq

gen dt = ehat_GARCH_sq - ehat_ARMA_sq
reg dt
ac dt
newey dt, lag(2)

//the coefficient of the dt is significant, meaning that arma model is significantly more
accurate than the GARCH model.
```