Quantitative Risk Management Project Report

Thomas Aujoux, Paul Wattellier

November 23, 2023

Part A: Stylized Facts and GARCH Modelling

This section focuses on the statistical analysis and modeling of the EURO STOXX 50 stock market index. The index measures the performance of the 50 largest and most liquid Eurozone stocks.

(0) Computing the Daily Returns

We will analyze the time series of asset prices S_t, specifically focusing on the "adjclose" column, which represents prices. To better understand the dynamics of asset prices, we calculate the daily log returns. The daily log return, denoted as r_t, is defined as:

$$r_t = \ln S_t - \ln S_{t-1} = \ln(1 + R_t).$$

We create two new pandas columns named "logreturn" to store the calculated log returns. Additionally, to express the values in percentage terms, we multiply these columns by 100. We can visualize the asset prices and daily returns with Figures 1 and 2.

Figure 1: Asset Prices Over Time

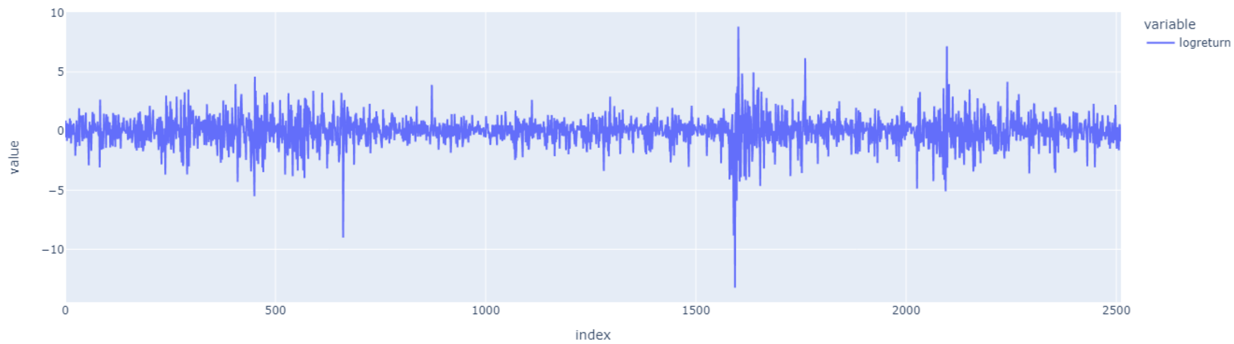

Figure 2: Daily Returns Over Time

(i) Descriptive Statistics and Distribution Analysis

In this section, we present a statistical analysis of the log-returns of asset prices.

The mean log-return is 1.1063%, and the standard deviation is 123.8184%. These values provide information about the central tendency and dispersion of the log returns.

Skewness measures the asymmetry in a data set, while kurtosis quantifies the "tailedness" of the distribution. For the log returns of SX5E, the empirical skewness is -0.7954, deviating from the theoretical skewness of 0 for a normal distribution. The negative skewness implies a potential prevalence of more extreme negative returns than positive returns. The empirical kurtosis is 10.2873, significantly higher than the theoretical kurtosis of 3 for a normal distribution. This higher kurtosis indicates heavy tails and more extreme values in the distribution compared to a normal distribution and suggests an increased likelihood of extreme events.

To further explore the distribution of log returns, we attempt to fit a normal distribution $N(\mu, \sigma)$ using the sample mean (μ) and standard deviation (σ) of the daily log returns. Figure 3 displays the histogram and kernel density estimate of log returns. The fitted normal distribution density is overlaid on the histogram.

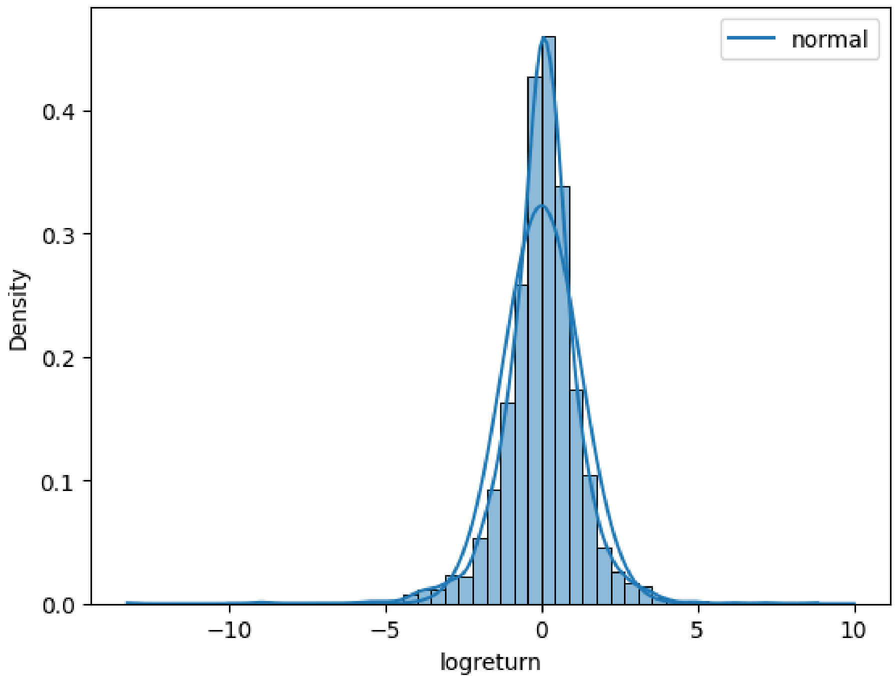

Figure 3: Histogram and Kernel Density Estimate of Log Returns

The fitted kernel estimator exhibits a more acute peak and heavier tails relative to the normal probability density function (pdf) as we said before. The histograms in Figure 3 illustrate that the empirical distribution of the SX5E log returns deviates from a Gaussian distribution.

(ii) Autocorrelation Function (ACF) Analysis

The autocorrelation function (ACF) plot for the daily return, denoted as $\rho(k) = \text{Cor}(r_t, r_{t-k})$ for $k = 0, 1, \ldots, 50$, is presented in Figure 4. The correlation values are negligible beyond $k = 0$, indicating that forecasting daily returns based on historical values may be challenging due to the lack of (linear) dependence.

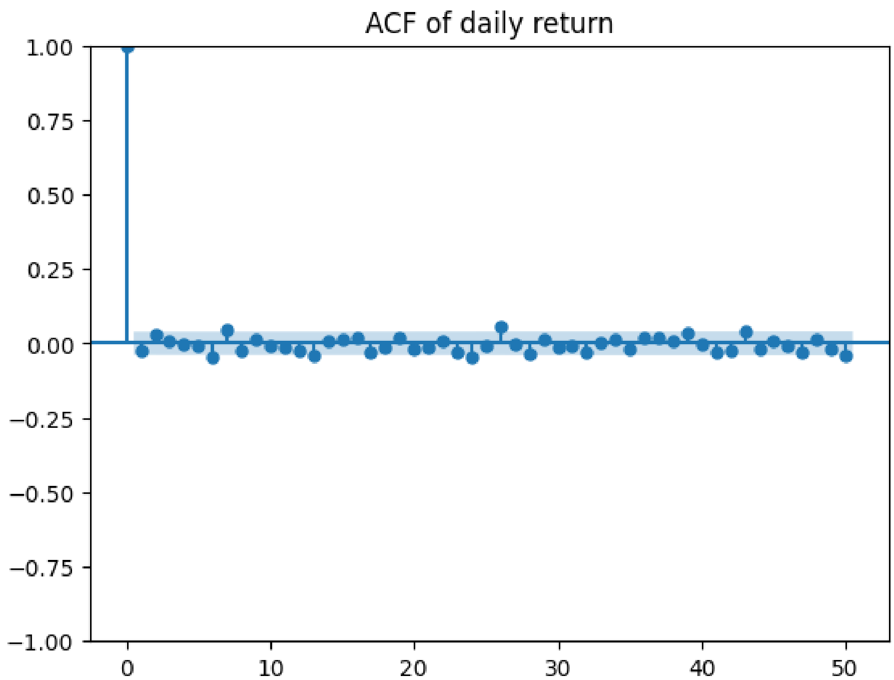

Figure 4: ACF Plot of Daily Returns

However, when considering the ACF of the absolute daily return, denoted as $\rho(k) = \mathrm{Cor}(|r_t|, |r_{t-k}|)$, the correlation values become more significant. This phenomenon is known as "volatility clustering," indicating that large price movements tend to be followed by more large movements in the near future. The persistence of significant correlation values for larger k suggests that the impact of a large move today can influence returns for an extended period.

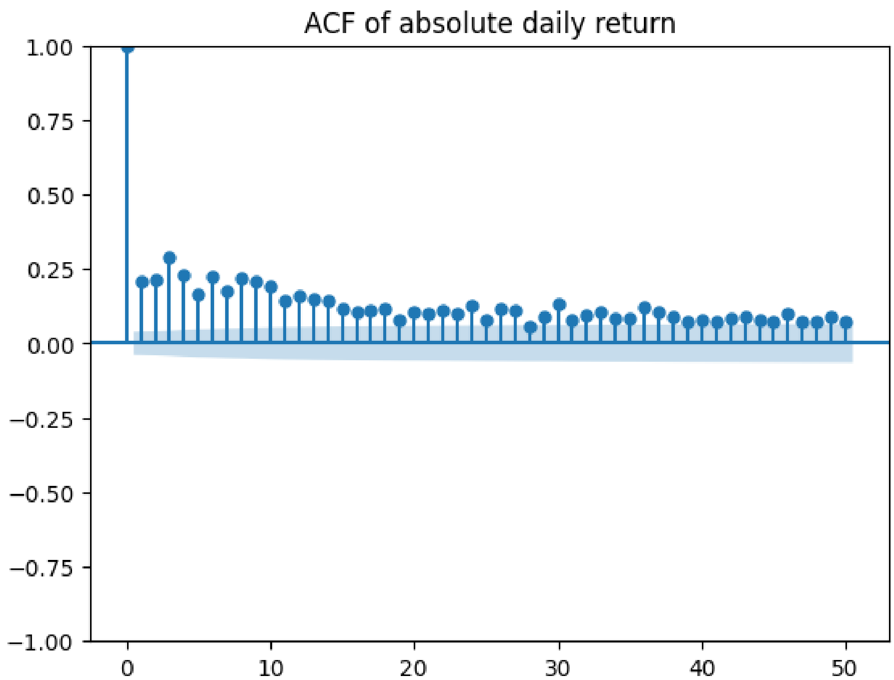

Figure 5: ACF Plot of absolute Daily Returns

(iii) GARCH(1,1) Model Fitting

In this section, we aim to fit a GARCH(1,1) model to the time series $X = (X_t)_t$, where X_t represents the daily log-return. The specification of the GARCH(1,1) model is given by:

$$X_t = \sigma_t Z_t, \tag{1}$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \tag{2}$$

where $(Z_t)_t$ is assumed to be normally distributed strict white noise with a common probability density function (pdf) given by:

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}.$$

To fit the GARCH(1,1) model, we maximize the log-likelihood function:

$$\ell(\alpha_0, \alpha_1, \beta_1; X) = \sum_{t=1}^{T} \left[-\frac{1}{2} \log \sigma_t^2 - \frac{X_t^2}{2\sigma_t^2} - \frac{1}{2} \log(2\pi) \right], \tag{3}$$

with respect to the GARCH parameters $(\alpha_0, \alpha_1, \beta_1)$ and the observed data $X = (X_t)_{t=0}^{T}$.

The maximization problem is formulated as:

$$\sup_{\alpha_0,\alpha_1,\beta_1} \ell(\alpha_0, \alpha_1, \beta_1; X), \tag{4}$$

subject to the constraints $\alpha_0 \geq 0$, $\alpha_1 \geq 0$, $\beta_1 \geq 0$, and the stationary condition $\alpha_1 + \beta_1 < 1$. We can use Python package called Arch which will provide us a solution for this problem.

Results

The constant mean (μ) is estimated to be 0.0476 with a t-statistic of 2.353, indicating statistical significance at a 5% level. This represents the average return in the model.

The GARCH(1,1) model includes ARCH and GARCH coefficients. The alpha (ARCH) coefficient is 0.1381, and the beta (GARCH) coefficient is 0.8251. Both coefficients are statistically significant.

The R-squared and adjusted R-squared values are both 0.000, suggesting that the model does not explain much of the variation in the data. The log-likelihood is -3747.84, and the AIC and BIC values are provided for model comparison, we will use them for the last question.

Results Interpretation

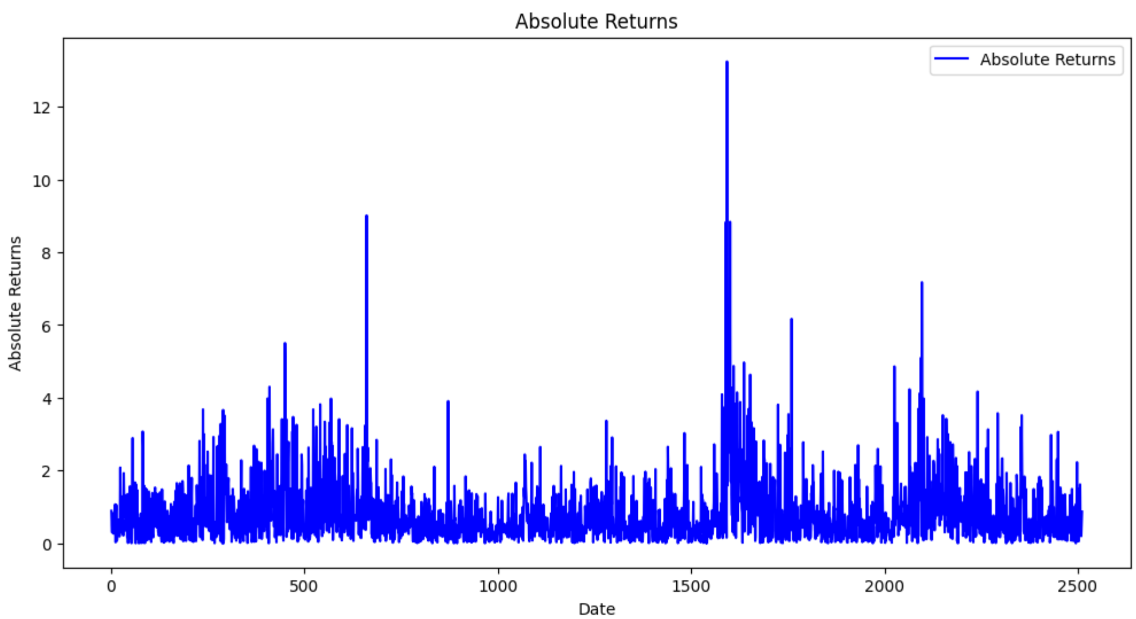

Figure 6: Plot of Absolute Returns

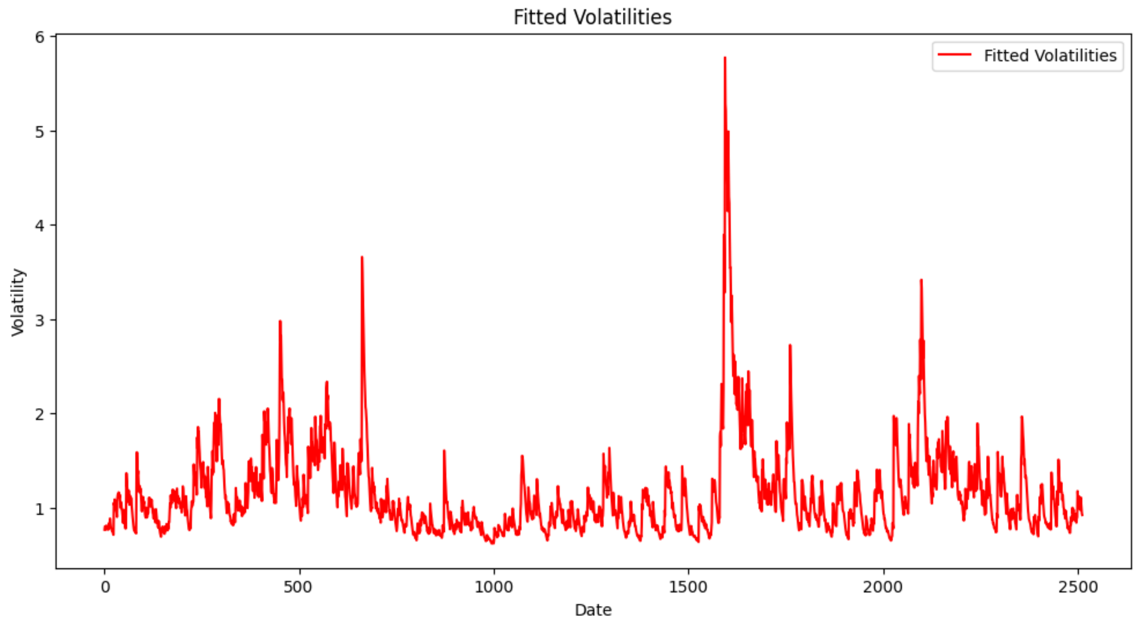

Figure 7: Plot of Fitted Absolute Returns

Figure 6 shows the magnitude of the log returns over time. Peaks in absolute returns may indicate periods of high volatility or large price movements.

Figure 7 shows fitted absolute returns estimated by the GARCH(1,1) model over time. It represents the model's attempt to capture the changing volatility patterns in the data. The model seems to perform well, the fitted volatilities seem to capture the volatility spikes seen in the absolute returns plot.

(iv) Goodness of Fit Assessment

To assess the goodness of fit of the GARCH(1,1) model estimated in part (iii), we can examine the standardized residuals and check if they look like an independent and identically distributed (i.i.d.) sample from a standard normal distribution using a statistical test. The standardized residuals are obtained by dividing the residuals by the conditional volatility.

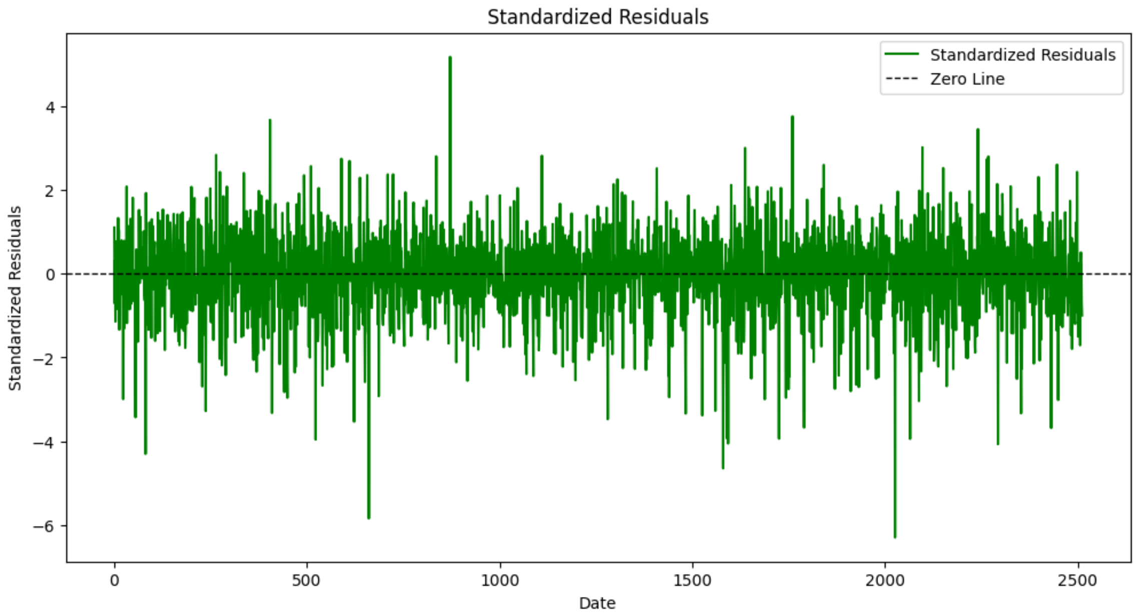

Figure 8: Plot of Standardized Residuals

To formally test whether the standardized residuals are independent and identically distributed from a standard normal distribution, we can use the Shapiro-Wilk test for normality. If the p-value is greater than the significance level (commonly 0.05), you fail to reject the null hypothesis, suggesting that the standardized residuals may be reasonably assumed to be normally distributed.

```
Shapiro-Wilk Test for Normality:

Test Statistic: 0.97
P-value: 5.16e-22
```

The results of the Shapiro-Wilk test for normality suggest that the model may not provide a good fit to the data. We can try other models to improve the results and to see if the standardized residuals are normally distributed and iid.

Conclusions

The GARCH(1,1) model has been successfully fitted to the log return data. The estimated coefficients for the mean, ARCH, and GARCH components are statistically significant. However, the model does not explain much of the variation in the data, as indicated by the low R-squared values. Additionally, the Shapiro-Wilk test suggests that the standardized residuals do not follow a normal distribution.

(v) ARMA(1,1)–GARCH(1,1) Model with Student t-distributed Innovations

We can repeat the same steps this time using a Student t-distribution innovation and adding the ARMA(1,1) part to the model. We can then compare the results with the model used in question (iii).

Results and comparaison in term of GARCH(1,1) and ARMA(1,1)-GARCH(1,1)

The coefficient for the ARMA(1,1) part of the model is estimated to be 0.919 for the AR(1) with a t-statistic of 12.296 and -0.94 for the MA(1) with a t-statistic of -14.603, so the coefficients of the ARMA(1,1) are statistically significant. This suggests that the ARMA(1,1) component is necessary and improve the fit of the model.

The ARMA(1,1)-GARCH(1,1) model includes ARCH and GARCH coefficients. The alpha (ARCH) coefficient is 0.147, and the beta (GARCH) coefficient is 0.835. Both coefficients are statistically significant.

The R-squared value is 0.00, suggesting that the model does not explain much of the variation in the data. The log-likelihood is -3644.24, and the AIC and BIC values are 7302.48 and 7343.28 whereas the AIC and BIC values for the model of the question (iii) where 7503.69 and 7527.00, this indicates an improvement in the model. So, we can conclude that the new specification improves the fit, compared to the model in (iii), including the ARMA component in the mean equation is necessary for capturing additional patterns in the data.

Goodness of Fit Assessment in terms of innovations

```
Kolmogorov-Smirnov test to test the T-student distribution:

Test Statistic: 0.022
P-value: 0.15
```

The results of the Kolmogorov-Smirnov test suggest that the model provides a better fit to the data compared to the model used in question (iii) because the p-value for the Kolmogorov-Smirnov test more than 0.05. So the Student t-distributed innovations are a better fit.

Conclusions

The ARMA(1,1)–GARCH(1,1) model with Student t-distributed innovations appears to provide a better fit to the data compared to the GARCH(1,1) model, based on the lower AIC value and a greater p-value for the test on the innovation. Moreover, the significance of the ARMA(1,1) coefficients demonstrates the necessity of the ARMA component.

Part B: Risk Measures

Introduction

In this section of the report, we delve into the practical application of quantitative risk management techniques, specifically focusing on the ProShares Short Dow30 (DOG) exchange-traded fund (ETF).

Our objective is to assist an investor, referred to here as Bob, who has chosen to invest in DOG as a means to short US stocks. Bob's investment strategy is underpinned by a desire to capitalize on potential declines in the Dow Jones index. However, given the inherent risks associated with such a strategy, especially in the volatile domain of stock markets, it is crucial to employ robust risk management techniques.

This section of the report is dedicated to implementing and evaluating one-day-ahead Value-at-Risk (VaR) and Expected Shortfall (ES) forecasts for Bob's investment. The analysis spans from an appropriately chosen starting date until 27 October 2023, offering a comprehensive view of the risk profile over a significant period.

To achieve a thorough risk assessment, we employ three distinct methodologies:

- **Historical Simulation (HS)**: This method utilizes a rolling window of past data to simulate future outcomes, offering a straightforward yet effective means of risk estimation.
- **Filtered Historical Simulation (FHS) with Exponentially Weighted Moving Average (EWMA)**: This approach enhances the traditional HS method by incorporating time-varying volatility, thereby offering a more nuanced view of risk that accounts for market dynamics.
- **FHS with GARCH Model**: This method combines the GARCH model's ability to capture time-varying volatility with the HS approach, providing a sophisticated tool for risk estimation.

Each of these methods has its merits and limitations, which will be explored in the subsequent sections. The report not only presents the results of these risk assessment methods but also includes a backtesting analysis to evaluate their performance. This comprehensive approach ensures that the recommendations made to Bob are grounded in empirical evidence and robust statistical analysis.

Historical Simulation (HS)

Methodology

In the Historical Simulation (HS) approach, as outlined in Part 2 of the Quantitative Risk Management course (slides 54-55), historical data is used to empirically estimate the distribution of future returns. This method presupposes that past financial data can be indicative of future risks. Our calculation of one-day-ahead Value-at-Risk (VaR) and Expected Shortfall (ES) forecasts involved the following steps:

- Data Preparation: We utilized daily adjusted closing prices of the ProShares Short Dow30 (DOG) ETF, covering the period from 30 October 2013 to 27 October 2023.

- Log Return Calculation: Daily log returns, r_t, were computed using the formula:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right)$$

  where P_t represents the adjusted closing price on day t.
- Rolling Window Scheme: Consistent with the course recommendations, we employed a rolling window of 500 trading days for each forecast.
- VaR and ES Calculation:
  - Value-at-Risk (VaR): VaR at a given confidence level α was estimated as the empirical percentile of the log returns within each 500-day window. Mathematically, the estimator for a one-day-ahead 95% VaR is given by:

$$\mathrm{VaR}_\alpha = -\inf\{x \in R : F(x) > 1-\alpha\}$$

    where F is the empirical distribution function of the log returns.
  - Expected Shortfall (ES): ES, representing the average of losses exceeding the VaR, was calculated as follows:

$$\mathrm{ES}_\alpha = -\frac{1}{n(1-\alpha)}\sum_{i=1}^{n} r_{(i)} I_{\{r_{(i)} \le \mathrm{VaR}_\alpha\}}$$

    where $r_{(i)}$ are the ordered log returns, and I is the indicator function.

Results

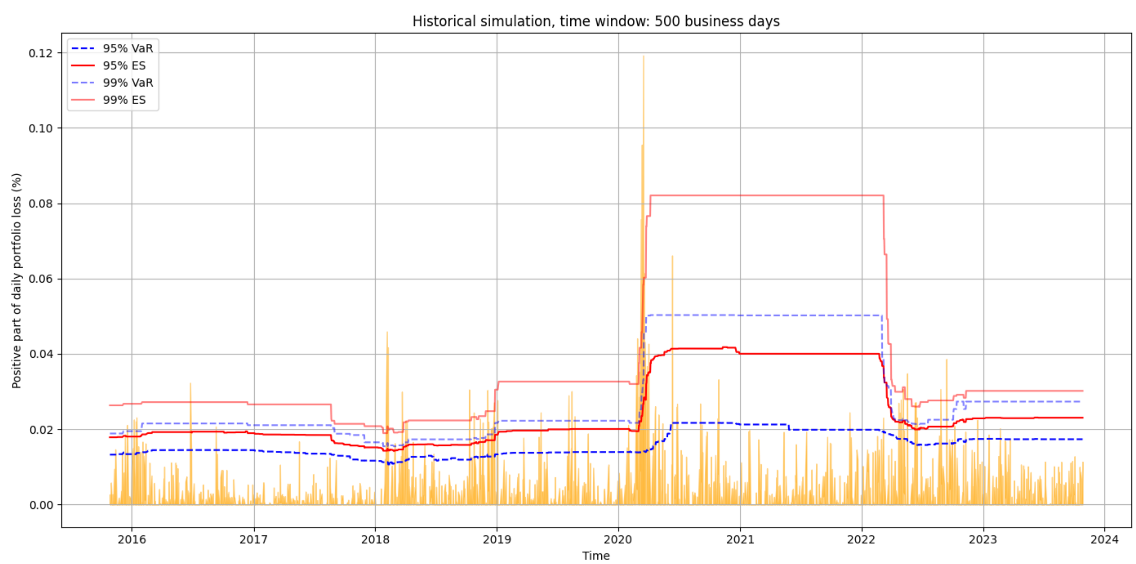

Figure 9: Historical Simulation

These values indicate the maximum expected loss (VaR) and the average of more extreme losses (ES) at the respective confidence levels, under normal market conditions.

Filtered Historical Simulation (FHS) with EWMA

In this part of the analysis, we apply the Filtered Historical Simulation (FHS) method enhanced with the Exponentially Weighted Moving Average (EWMA) model for volatility forecasting. This approach, as referenced in Part 3 of the Quantitative Risk Management course (slide 97), is particularly adept at capturing the dynamic nature of financial market volatility.

EWMA Volatility Forecast:
The EWMA model is employed to estimate time-varying volatility of the log returns, r_t, of the investment. In the EWMA model, the volatility forecast for time t, σ_t^2, is given by the formula:

$$\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1 - \lambda) r_{t-1}^2$$

where r_{t-1} represents the log return at time $t-1$. For the purpose of this analysis, we make the simplifying assumption that $r_t = X_t - \mu_t$ with $\mu_t = 0$, as suggested in the coursework instructions. Thus, r_t simplifies to the log return X_t. The decay factor λ is set to 0.94, given $\alpha = 0.06$.

Standardized Residuals:
After computing the EWMA volatility, we derive the standardized residuals, z_t, using:

$$z_t = \frac{r_t}{\sigma_t}$$

Historical Simulation on Standardized Residuals:
Utilizing a rolling window of 500 trading days, we then apply historical simulation to these standardized residuals to estimate VaR and ES.

VaR and ES Calculation: The VaR and ES for the standardized residuals are computed using the following expressions:

- Value-at-Risk (VaR): The VaR at confidence level α is estimated by:

$$\text{VaR}_\alpha^z = -\inf \{x \in R : F_z(x) > 1 - \alpha\}$$

  where F_z is the empirical distribution function of the standardized residuals.

- Expected Shortfall (ES): The ES is determined as the mean of the residuals that exceed the VaR as previously :

$$\text{ES}_\alpha^z = -\frac{1}{n(1-\alpha)} \sum_{i=1}^{n} z_{(i)} I_{\{z_{(i)} \leq \text{VaR}_\alpha^z\}}$$

Results

The FHS with EWMA forecasts for the initial days of the analysis period are as follows:

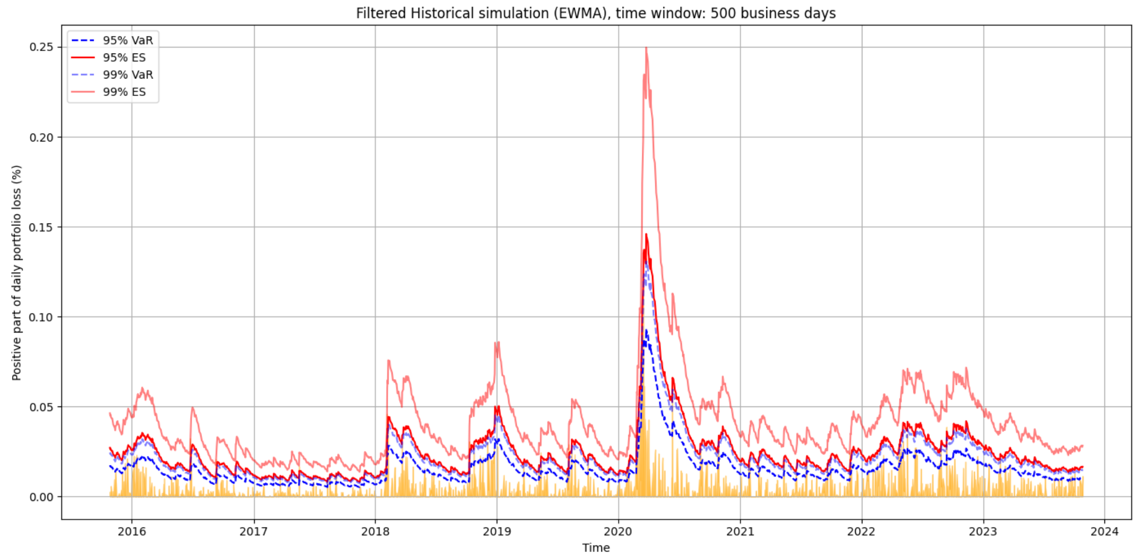

Figure 10: Filtered historical Simulation (with EWMA)

These forecasts account for the dynamic nature of market volatility, offering a nuanced view of the risk profile of the investment.

Filtered Historical Simulation (FHS) with GARCH

In Method (iii), we implemented the Filtered Historical Simulation (FHS) method combined with a GARCH(1,1) model, a sophisticated approach designed to capture the time-varying nature of volatility in financial markets.

- GARCH(1,1) Model Estimation: We estimated a GARCH(1,1) model on the log returns data. In this model, the conditional variance, σ_t^2, is expressed as:

$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2$$

  Here, r_{t-1} are the lagged log returns, and σ_{t-1}^2 is the lagged conditional variance.

- Standardized Residuals Calculation: From the GARCH model, we calculated the standardized residuals, z_t, which normalize the log returns by the conditional standard deviation:

$$z_t = \frac{r_t}{\sigma_t}$$

- Historical Simulation on Standardized Residuals: Applying historical simulation to these standardized residuals, we used a rolling window of 500 trading days to estimate VaR and ES, reflecting the dynamically changing risk profile.

- The Value-at-Risk (VaR) and Expected Shortfall (ES) are computed based on the empirical distribution of these standardized residuals. The results include the VaR and ES forecasts at different confidence levels, such as 95% and 99%.

Results

Our implementation yielded the following forecasts for the initial days of the analysis period:

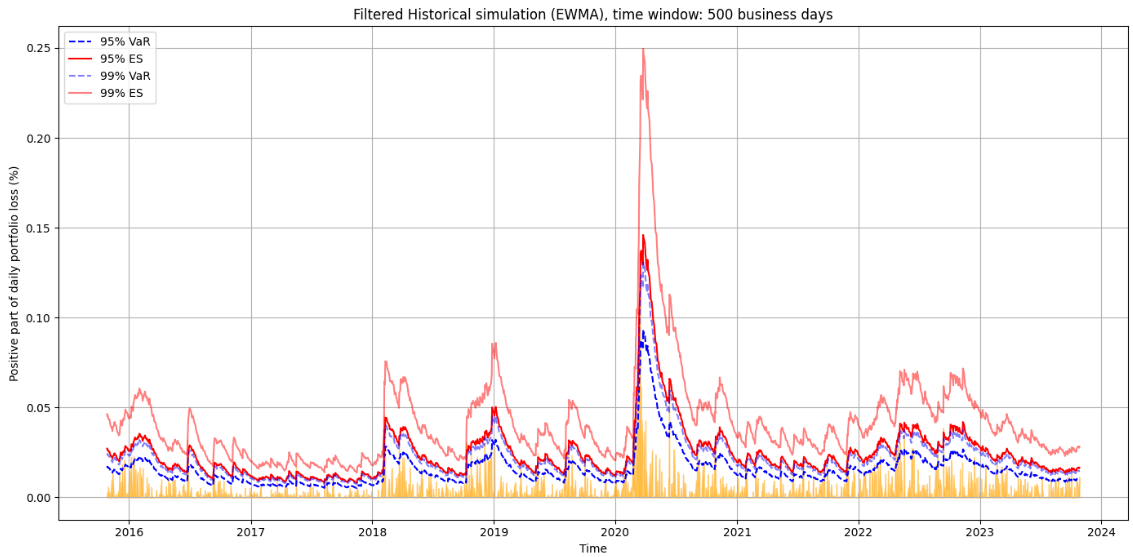

Figure 11: Filtered historical Simulation (with GARCH)

These results reflect the potential losses, adjusted for the changing volatility in the market, and provide a nuanced view of the investment's risk profile.

Unconditional Coverage Test (Kupiec's Test)

The Unconditional Coverage Test is used to assess whether the observed frequency of Value-at-Risk (VaR) violations aligns with the expected frequency under the specified confidence level. The null hypothesis for this test is that the VaR model correctly predicts the proportion of violations.

Given a VaR model at the α confidence level, the number of observed violations should theoretically be close to $(1-\alpha)$ percent of the total observations. The test statistic for Kupiec's Test is computed as:

$$LR_{uc} = -2\ln\left((1-\alpha)^{N-n}\cdot\alpha^{n}\right) + 2\ln\left(\left(\frac{N-n}{N}\right)^{N-n}\cdot\left(\frac{n}{N}\right)^{n}\right)$$

where N is the total number of observations, n is the number of observed violations, and α is the confidence level of the VaR model. Under the null hypothesis, the test statistic follows a chi-square distribution with one degree of freedom. A low p-value (e.g., less than 0.05) indicates that we can reject the null hypothesis, suggesting that the VaR model does not correctly predict the proportion of violations.

The results for our previous forecastings are the followings :

The table presents the results of the Unconditional Coverage Test (Kupiec Test) for VaR forecasts at different confidence levels (95% and 99%) and using different methods (HS, EWMA,

Method	LRuc	P-value	Violations
VaR 95% HS	10.338	0.0013	99.0
VaR 95% EWMA	16.274	0.0001	98.0
VaR 95% GARCH	10.338	0.0013	99.0
VaR 99% HS	-86.296	1.0000	30.0
VaR 99% EWMA	-34.682	1.0000	24.0
VaR 99% GARCH	1.380	0.2402	20.0

Table 1: Kupiec's Test

GARCH). The "LRuc" column represents the test statistic, the "P-value" column indicates the p-value of the test, and the "Violations" column shows the number of violations observed in the data.

At the 95% confidence level, all three methods (HS, EWMA, GARCH) have test statistics that exceed the critical value, resulting in p-values below the significance level of 0.05. This implies that the null hypothesis (H0) that the VaR models are correctly specified is rejected. Additionally, there is a high number of violations (above 95) for all methods at this confidence level, indicating that the models consistently underestimated risk.

At the 99% confidence level, the situation changes. The HS method exhibits a highly negative test statistic and a p-value of 1, suggesting that it overestimated risk at this level, and the number of violations decreases to 30. The EWMA method also overestimates risk, but to a lesser extent, with a p-value of 1 and 24 violations. In contrast, the GARCH method shows a positive test statistic with a p-value of 0.24, indicating that it may provide a more accurate estimate of risk at the 99% confidence level, with only 20 violations.

In summary, the coursework results suggest that, while all methods perform poorly at the 95% confidence level, the GARCH method shows promise in providing more accurate VaR estimates at the 99% confidence level, as it has a positive test statistic and a relatively higher p-value compared to the other methods. Further analysis and validation may be necessary to determine the most suitable method for managing risk in Bob's investment.

Independence Test

The Independence Test, also known as the Christoffersen Test, assesses whether the violations of Value at Risk (VaR) are independent over time. It is a critical test in risk management because it evaluates whether the risk model accurately captures the temporal dependencies in financial returns. The test is based on the assumption that if VaR violations are not independent, it can lead to increased exposure to risk.

In the context of the coursework, the Independence Test results for different VaR forecasting methods and confidence levels are summarized in the table below:

Interpretation:

The Independence Test assesses the temporal dependence of VaR violations for various forecasting methods and confidence levels. Key findings are as follows:

HS Method: At both 95% and 99% confidence levels, the test suggests significant departure from independence, with low p-values, indicating a potential issue with temporal dependence.

Method	Independence Test 95%	Independence Test 99%
HS	Chi-Square Statistic: 5.989749 P-value: 0.014389	Chi-Square Statistic: 5.562463 P-value: 0.018350
FHS EWMA	Chi-Square Statistic: 0.012405 P-value: 0.911318	Chi-Square Statistic: 0.292938 P-value: 0.588344
FHS GARCH	Chi-Square Statistic: 5.382602 P-value: 0.020339	Chi-Square Statistic: 0.202614 P-value: 0.652619

Table 2: Independence Test Results for the Different Methods and Confidence Levels

Filtered Historical Simulation EWMA Method: The results show a higher degree of independence compared to the HS method, with Chi-Square statistics close to zero and high p-values at both confidence levels, suggesting less temporal dependence.

Filtered Historical Simulation GARCH Method: At the 95% confidence level, there is a significant departure from independence, but the significance is lower at the 99% level, indicating that temporal dependence may be weaker.

These findings emphasize the importance of choosing an appropriate forecasting method, as it can impact the independence of VaR violations. Risk managers should consider these results when selecting a method to ensure the accuracy of risk assessments.

Joint Tests

In the context of the coursework, the Joint Test for Coverage and Independence results for different VaR forecasting methods and confidence levels are summarized in the table below:

Method	Joint Test 95%	Joint Test 99%
HS	Joint Test Statistic: 16.327459 P-value: 0.014389	Joint Test Statistic: -80.733096 P-value: 1.000000
FHS EWMA	Joint Test Statistic: 16.286522 P-value: 0.911318	Joint Test Statistic: -34.389272 P-value: 1.000000
FHS GARCH	Joint Test Statistic: 15.720313 P-value: 0.020339	Joint Test Statistic: 1.582280 P-value: 0.652619

Table 3: Joint Test for Coverage and Independence Results for Different Methods and Confidence Levels

These results shed light on the performance of various VaR forecasting methods:

- **HS Method**: At a 95% confidence level, the Joint Test Statistic suggests significant issues with coverage accuracy and/or temporal dependence, supported by a low p-value. However, at a 99% confidence level, the observed issues may not hold due to a high p-value.

- **FHS EWMA Method**: For both confidence levels, the Joint Test Statistic is positive, and the p-values are high, indicating that this method does not exhibit significant issues with coverage accuracy or temporal dependence.

- **FHS GARCH Method**: At the 95% confidence level, there are potential concerns with coverage accuracy and/or temporal dependence, but the significance is moderate. At the 99% confidence level, these concerns are less significant based on the higher p-value.

These findings underscore the critical role of choosing an appropriate VaR forecasting method, considering both coverage accuracy and temporal dependence, to ensure accurate risk assessments in financial management.

Conclusion

The backtesting results of the different Value-at-Risk (VaR) forecasting methods for Bob's investment, as shown in Table 1, provide insight into each method's accuracy in estimating potential losses. At the 95% confidence level, the Historical Simulation (HS), Exponentially Weighted Moving Average (EWMA), and GARCH models all reject the null hypothesis that the VaR models are correctly specified, evidenced by their low p-values below the 0.05 threshold and the high number of violations. This suggests these models consistently underestimated risk. However, at the 99% confidence level, the GARCH model stands out with a positive test statistic and a p-value of 0.24, implying a better fit as it has only 20 violations, which is close to the expected number. In contrast, the HS and EWMA methods display a large deviation from expected violations, indicated by their highly negative test statistics and p-values of 1.00, suggesting they overestimated risk. Thus, for Bob's investment, the GARCH method appears to be the most suitable for risk management at the 99% confidence level, offering a more accurate reflection of the actual risk when compared to the other methods.