



Time Series Analysis on Qualcomm Inc..

FA 542: Time Series with Applications to Finance
Final Project

Written By:
Andre Sealy

Federica Malaisura

Swapnil Pant

Contents

1	Asset Overview	2
1.1	Properties of the Time-Series	2
1.2	Descriptive statistics	2
2	Visualization	6
2.1	Unit-root and Seasonality	8
3	ARIMA Model	12
3.1	Model Selection	14
3.2	Forecasting	15
4	Vector Autoregressive Models	18
4.1	VAR Analysis	18
4.2	Cointegration	20
5	ARCH/GARCH Volatility Models	22
5.1	ARCH	22
5.2	GARCH	24
5.2.1	GARCH(1,1)	24
5.3	Forecasting	26
A	Figures	29
B	Tables	30

Chapter 1

Asset Overview

The asset class for our time series analysis consists of equities primarily in the Technology sector. The reason for choosing the technology sector out of many industry classifications is multifaceted and involves numerous foundational considerations. First, we consider that tech stocks often have higher betas (meaning, they move with the market) on average, suggesting higher annualized volatility. Second, they are more sensitive to new information related to innovation cycles, such as new product releases, software updates, and research. Finally, we have the growth aspect, which involves the shifts in macroeconomic trends (such as interest rates), global demand, and supply chains. The border macroeconomic environment and the fundamentals make tech stocks an ideal candidate for time series analysis.

1.1 Properties of the Time-Series

In this section, we will outline foundational descriptive statistics, which involve the statistical moments and distributions of our sample securities and time series analysis.

1.2 Descriptive statistics

First, we outline a few sample statistics for the stocks in our sample. These statistics involves the sample mean, denoted by

$$\hat{\mu}_x = \frac{1}{T} \sum_{t=1}^T x_t, \tag{1.1}$$

the sample standard deviation, denoted by,

$$\hat{\sigma}_x = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (x_t - \hat{\mu}_x)^2}, \quad (1.2)$$

the sample skewness, denoted by

$$\hat{S}(x) = \frac{1}{(T-1) \hat{\sigma}_x^3} \sum_{t=1}^T (x_t - \hat{\mu}_x)^3, \quad (1.3)$$

and the sample kurtosis (excess kurtosis), denoted by

$$\hat{K}(x) = \frac{1}{(T-1) \hat{\sigma}_x^4} \sum_{t=1}^T (x_t - \hat{\mu}_x)^4. \quad (1.4)$$

The sample mean, $\hat{\mu}_x$ represents the daily average simple (or log) returns, where T are the number of days in the sample. The sample standard deviation, denoted by $\hat{\sigma}_x$, represents the daily realized volatility, the universal risk measurement. Table (1.1) shows the descriptive statistics of the tech stocks in our sample of the daily sample returns. The start date of our analysis is January 3rd, 2007, which gives us a time interval of roughly 18 years. In addition to the sample mean, standard deviation, skewness, and excess kurtosis, we have also decided to include the minimum and maximum prices within this time period for a better perspective.

Table 1.1: Descriptive statistics of Tech Stocks (*Daily Simple Returns %*)

Security	Mean ($\hat{\mu}_x$)	Std Dev ($\hat{\sigma}_x$)	Skewness ($\hat{S}(x)$)	Kurtosis ($\hat{K}(x)$)	Min	Max
QCOM	0.05	2.17	0.27	9.53	-15.25	23.20

The returns between our sample securities are roughly similar when we look at daily log returns in Table (1.2).

The sample skewness measures the asymmetry of a probability distribution around the sample mean. It helps determine whether our sample is skewed more towards one side of the distribution. If the sample skewness, $\hat{S}(x) \approx 0$, the distribution is symmetrical. Based on skewness, QCOM, CSCO is roughly symmetrical, as the skewness is approximately 0. On the hand, the sample kurtosis measures how heavy the tails of the probability distribution are, which helps assess the presence of extreme values and the frequency of their occurrence. A normal distribution has a sample kurtosis of $\hat{K}(x) \approx 3$, which indicates a normal distribution. All of the securities in our sample have extremely high kurtosis, which is natural for equities,

Table 1.2: Descriptive statistics of Tech Stocks (*Daily Log Returns %*)

Security	Mean ($\hat{\mu}_x$)	Std Dev ($\hat{\sigma}_x$)	Skewness ($\hat{S}(x)$)	Kurtosis ($\hat{K}(x)$)	Min	Max
QCOM	0.21	1.06	-1.53	5.03	-16.54	3.18

considering financial markets experience extreme price movements more frequently than a normal distribution would predict. Financial markets also show periods of high and low volatility; when volatility spikes, returns tend to have more extreme values.

Considering the sample skewness, AMD is much further away from 0. There is some slight positive skew QCOM. We do not have a single sample statistic that we can use to differentiate skewed versus unskewed. However, we can test for skewness using the Jarque and Bera test for normality.

The Jarque-Bera (JB) test is a statistical test used to check whether a dataset follows a normal distribution. It does this by examining the skewness and kurtosis of the dataset with the following formula:

$$JB = \underbrace{\frac{\hat{S}^2(r)}{\sqrt{6/T}}}_{skewness} + \underbrace{\frac{[\hat{K}(r) - 3]^2}{24/T}}_{kurtosis} \quad (1.5)$$

which is distributed as a chi-square, χ^2 , random variable with 2 degrees of freedom. Given the sample of returns, $\{r_1, \dots, r_T\}$, to test the skewness of the returns, we consider the following hypothesis test

$$\begin{aligned} H_0 : S(r) &= 0 \\ H_a : S(r) &\neq 0 \end{aligned}$$

where the t -statistic is

$$t = \frac{\hat{S}(r)}{\sqrt{6/T}} \quad (1.6)$$

We would also need to conduct a hypothesis test for the excess kurtosis, which is

$$\begin{aligned} H_0 : K(r) - 3 &= 0 \\ H_a : K(r) - 3 &\neq 0 \end{aligned}$$

where the t -statistics is the following:

$$t = \frac{\hat{K}(r - 3)}{\sqrt{24/T}} \quad (1.7)$$

We provide the JB-test for statistical normality with the skewness, $\hat{S}(x)$, kurtosis, $\hat{K}(x)$, t-statistic, \mathcal{X}^2 , and p-values. We reject the null-hypothesis, H_0 of normality if the p -value of the JB statistic is less than the significance level. As the Table (1.3) shows, the p -values of the JB-test for our sample securities are significantly below the 5% significance level, which indicates that MSFT, QCOM, AMD are skewed to the right (positively skewed) and ADBE are negatively skewed.

Table 1.3: Jarque-Bera Statistical Test for Normality for log daily returns

Security	Skewness ($\hat{S}(x)$)	Kurtosis ($\hat{K}(x)$)	t -statistic (\mathcal{X}^2)	p -value
QCOM	0.27	9.53	14733.69	2.2e-16

Chapter 2

Visualization

We first begin our analysis by looking at cumulative returns, an essential part of time series analysis. Unlike daily or periodic returns, which can be volatile, cumulative returns smooth out fluctuations and show the overall trends of a time series. They also help compare the growth of our asset class over different time horizons and identify long-term trends in the market. Cumulative returns will serve as the basis for other predicting and forecasting techniques. Figure (2.1) shows the cumulative returns for QCOM. Cumulative returns ignore



Figure 2.1: Cumulative Returns for the sample securities

risk and measure the total percentage return, whereas metrics such as the standard deviation, $\hat{\sigma}_x$, and kurtosis, $\hat{K}(x)$, account for risk. The idea of studying volatility is that the returns of our asset $\{r_t\}$ have a serial correlation, or it may merely be stationary. The Autocorrelation Function (ACF) measures the correlation between a time series and its own past values at different lags. It helps identify patterns, dependencies, and stationarity in time series data.

When considering a weakly stationary return series, r_t , the linear dependence between r_t and past values of t_{t-i} is called the lag- ℓ autocorrelation of r_t , denoted by the following:

$$\hat{\rho}_\ell = \frac{\sum_{t=\ell+1}^T (r_t - \bar{r})(r_{t-\ell} - \bar{r})}{\sum_{t=1}^T (r_t - \bar{r})^2}, \quad 0 \leq \ell < T - 1 \quad (2.1)$$

For testing several autocorrelations jointly, we have the Ljung and Box test, denoted by,

$$\mathcal{Q}(m) = T(T+2) \sum_{\ell=1}^m \frac{\hat{\rho}_\ell^2}{T-\ell} \quad (2.2)$$

The hypothesis test for Ljung and Box is established as follows:

$$H_0 : \rho_1 = \dots = \rho_m = 0$$

$$H_a : \rho_i \neq 0$$

for some $i \in \{1, \dots, m\}$. We assume that $\{r_t\}$ is an iid sequence and $\mathcal{Q}^*(m)$ is a chi-squared random variable, \mathcal{X} , with m degrees of freedom. The decision rule is to reject H_0 if $\mathcal{Q}(m) > \mathcal{X}_\alpha^2$, where \mathcal{X}_α^2 denotes the 100(1 - α)th percentile of a chi-squared distribution.

Figure (??) shows the sample autocorrelation functions of the monthly log returns from January 2007 to November 2025. We have used a maximum of 20 lags. In each plot, two horizontal dashed lines denote two standard error limits of ACF. The sample ACFs or the MSFT plot are very close to each other, and they suggest that the serial correlations of monthly MSFT log stock returns are very small, if any. The sample ACFs are all within their two standard error limits, indicating that they are not significantly different from zero at the 5% level.

Regarding the log monthly returns for the other securities in our sample, we find there is at least one lag that determines a significant serial correlation with previous returns. For QCOM, this occurs at lag 17, From looking at the ACF plots, it would appear that these lags barely reach the rejection threshold of 2 times the standard error. However, we can perform the Ljung-Box statistic determine if there are some significant serial correlations at the 5% level for all return series. The Ljung-Box statistic with the max lag gives us a p-value of 0, so we reject the null-hypothesis.

As we can see from the test statistics and the p -values, the only securities with a significant serial correlation with monthly log returns are QCOM. What we can interpret from this data, at least for long monthly returns, is that there is no significant serial correlation. It may be worth using a more frequent observation series, such as daily.

We have also looked at the Partial Autocorrelation Function (PACF) of the log monthly

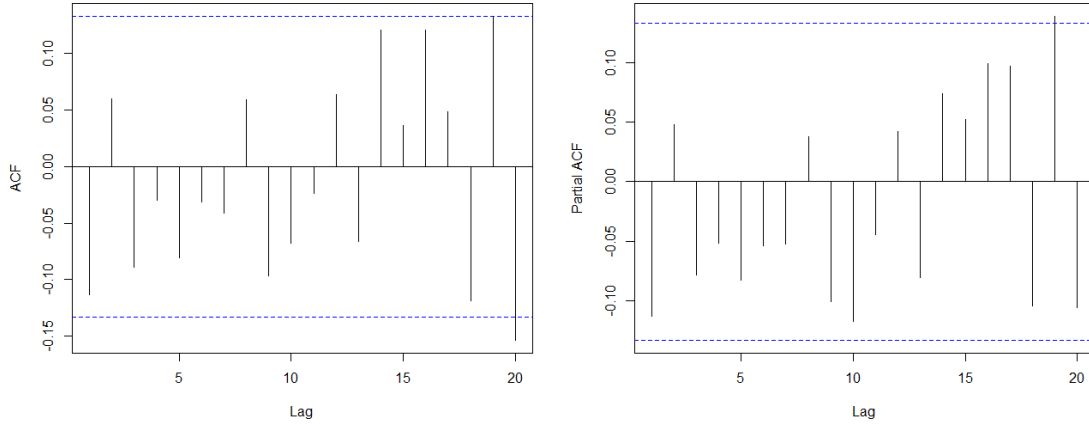


Figure 2.2: ACF and PACF of QCOM

returns. The PACF is used to directly measure the relationship between a time series and its past values at different lags while removing the influence of the intermediate lags. Unlike the ACF, which includes direct and indirect correlations, the PACF isolates the pure effect of each lag on the present value.

Figure (??) shows the PACF for the log monthly returns for our sample securities. We have used a maximum of 20 lags. After accounting for intermediate lags, each bar in the PACF plot represents the correlation between the time series and its lagged version. The dashed lines represent the 95% confidence interval. Any bar extending beyond these lines is statistically significant. ADBE and CSCO all appear to have significant lags.

ADBE have significant lags at month 6 and month 12; meaning the returns 6 and 12 months have a direct influence on the current month returns. This suggest that monthly log returns of ADBE may have some seasonal dependency. With CSCO, we have significant lags at month 6 and 14, suggesting the log monthly returns from 6 and 14 months ago have a direct influence on the log monthly returns in the current month.

2.1 Unit-root and Seasonality

To test whether the log price p_t of an asset follows a random walk, we would need to employ the following models

$$\begin{aligned} p_t &= \phi_1 p_{t-1} + \epsilon_t \\ p_t &= \phi_0 + \phi_1 p_{t-1} + \epsilon_t \end{aligned} \tag{2.3}$$

where ϵ_t denotes the error term. We need to consider the null hypothesis

$$\begin{aligned} H_0 : \phi_1 &= 1 \\ H_a : \phi_1 &< 1. \end{aligned} \tag{2.4}$$

Where the test statistic is

$$\hat{\phi}_1 = \frac{\sum_{t=1}^T p_{t-1} p_t}{\sum_{t=1}^T p_{t-1}^2}, \quad \hat{\sigma}_\epsilon^2 = \frac{\sum_{t=1}^T (p_t - \hat{\phi}_1 p_{t-1})^2}{T - 1} \tag{2.5}$$

where $p_0 = 0$ and T is the sample size. If the null hypothesis shows that $\phi_1 = 1$, we fail to reject the null hypothesis, which say that the time series has a Unit Root. Otherwise, we can reject the null hypothesis and consider the time series as stationary. We perform this hypothesis test with a technique commonly referred to as the augmented Dickey-Fuller (ADF) unit-root test.

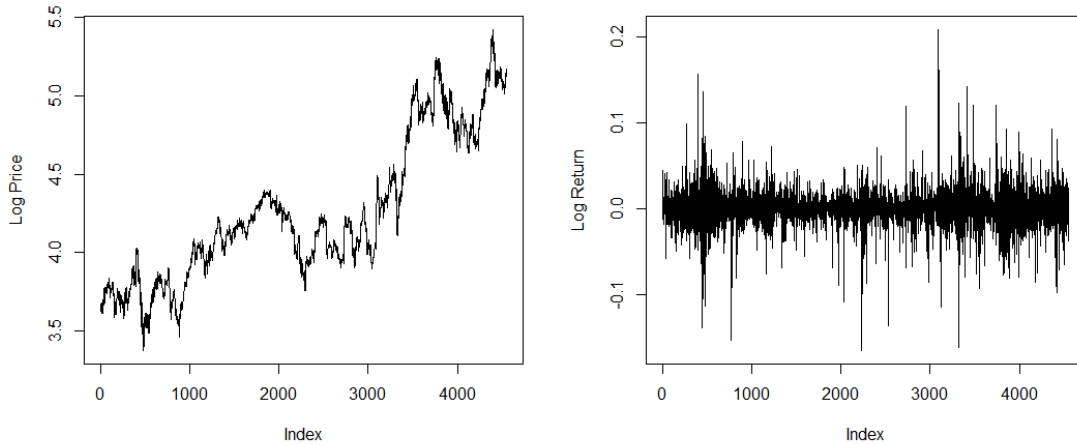


Figure 2.3: Time Series of QCOM Log Prices and Returns

Figure (2.3) shows the log returns and prices of the QCOM. We took the log transformation for two reasons. First, it is used to handle the exponential growth of the series, which the plot on the left confirms that the growth is linear on a log scale. Second, the transformation is used to stabilize the variability of the series. With these adjustments in mind, we can see that the log price exhibit an upward trend. On the other hand, the variability changes throughout the time series when it comes to the log returns. Despite this, the log of the log returns is clearly stationary. We implemented the augmented Dickey-Fuller test on both the log prices and returns. Table (2.1) shows the output of the augmented Dickey-Fuller test for the log prices and returns of QCOM. As we can see our suspicions were correct; we fail to

Table 2.1: Augmented Dick-Fuller Test for QCOM

Time Series	Lag order	t-statistic	p-value	Outcome
Log Prices	16	-3.051	0.1333	Unit-Root
Log Returns	16	-16.254	0.01	Stationary

reject the null-hypothesis for the log prices, but reject the null-hypothesis for the log returns. As such, the log returns are stationary, while the log prices exhibit unit-root.

We establish the time series' seasonality by examining the log prices' sample ACF. Figure (2.4) shows the ACF plots of QCOM with varying differencing and seasonality implementations. We have used a maximum lag of 100, with the time and seasonal differencing of once every 30 lags. The autocorrelation plot on the top right contains a very high and slow decay, starting from 1 at lag 0, which decreases gradually. This is typical of a non-stationary time series, where the values of the sample ACF are highly dependent on their past, such as a random walk.

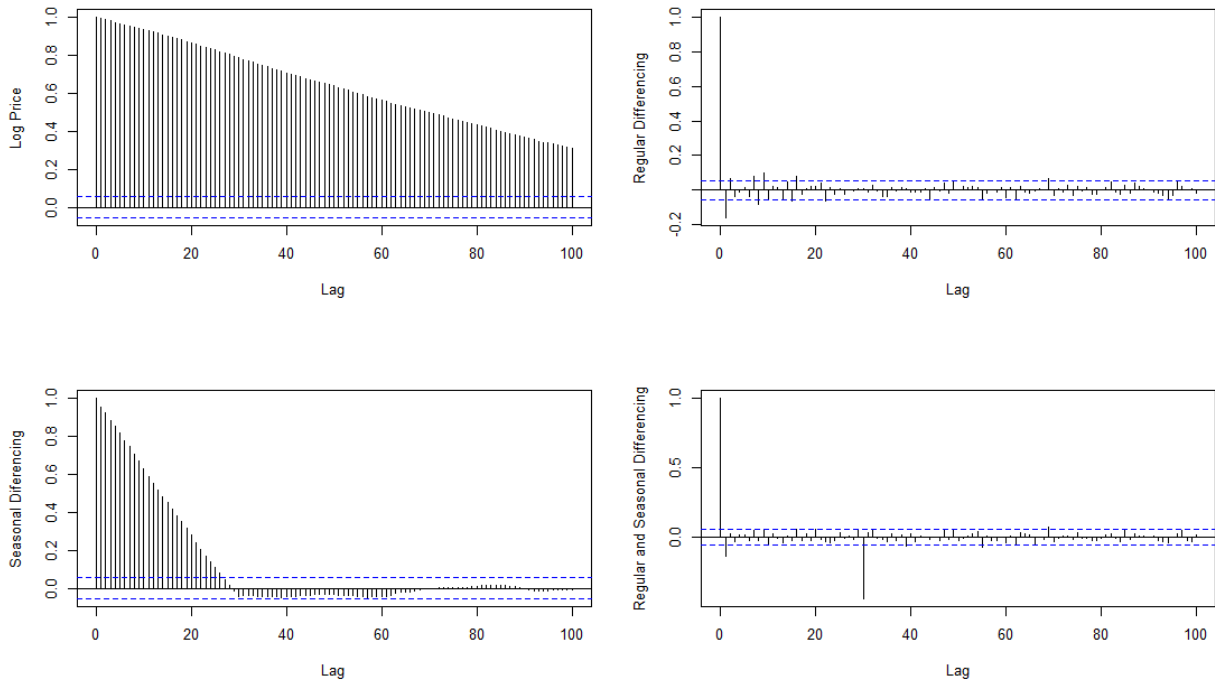


Figure 2.4: ACF Plot of QCOM

The plot on the top right shows the ACF of the First differenced log prices of the daily log prices. We can see a sharp drop in the autocorrelation at lag 1, where the other lags are within the confidence bounds. The first differencing has successfully removed the trend,

producing a series that now looks more like white noise.

On that bottom left plot, we have the ACF of the seasonal differencing, with a strong autocorrelation at lower lags, which slowly decreases to zero around lag 30. The differencing involves the following formula

$$\Delta_{30}(\Delta x_t) = (1 - B^{30})\Delta x_t = \Delta x_t - \Delta x_{t-30} \quad (2.6)$$

After lag 30, the values fluctuate randomly around zero. This suggests that seasonal differencing alone removes the long-run periodicity but not the trend of the time series. In other words, seasonal differencing alone is not enough to transform the plot into a stationary time series.

The final plot on the bottom right shows the ACF after regular and seasonal differencing. The pattern shows that almost all autocorrelations are within the confidence bands, except for a substantial spike at lag one and lag 30, which suggest a seasonal lag. As we can see, combining regular and seasonal differencing yields a stationary series, and the residual autocorrelations are minimal. These transformations show that the plots on the bottom left and right show a stationary transformation. Fitting an ARIMA model would involve implementing either regular differencing or both regular and seasonal differencing.

Chapter 3

ARIMA Model

The goal of this section is to fit log prices and log returns into an autoregressive (AR), moving average (MA), or some combination of the two (ARMA, ARIMA). Figure (3.1) shows the ACF for the log prices (a), PACF of the log prices (b), ACF for the log returns (c) and the PACF for the log returns (d). For AR models, we would expect the ACF to show a gradual decay where the tail cuts off at some point, where the PACF would sharply cut off at lag p , then drop to approximately zero. We expect the opposite for MA models, where the ACF significantly cuts off at lag q (then drops off to near zero), and the ACF gradually decays. On the other hand, the ACF and PACF for the ARMA(p,q) model will not clearly indicate where the lags cut off.

We first examine the log prices of QCOM. As we can see from Figure (3.1), the ACF plot (a) for the log prices shows a very sharp spike at lag 0, and a significant lag around point 30; the rest is essentially noise. This plot is not very useful, because the time series is non-stationary. We have the same pattern when looking at the PACF plot (b) for the log prices, as we see mostly noise and a large spike around lag 30. Again, this plot is not very reliable since stationary is required for the PACF/ACF patterns to inform model selection.

Looking at the ACF plot (c) log returns, we have a slight negative spike at lag 1. We also have a lot of significant lags that show up every 3 or 4 lags. Looking at the PACF plot (d), there are many small spikes, some of which are at lag 5, 8, 15, etc., but none are dominant. There is no clear cut-off in the PACF plot; we also do not see a gradual decline, which is consistent with the MA process. Using this plot, we cannot determine whether or not we should use an MA or AR model. We also consider using ARMA(p,q) models.

The ARMA model (AutoRegression Moving Average) is a time series model used to describe stationary processes, which combines the past values of the AR(p) model and the

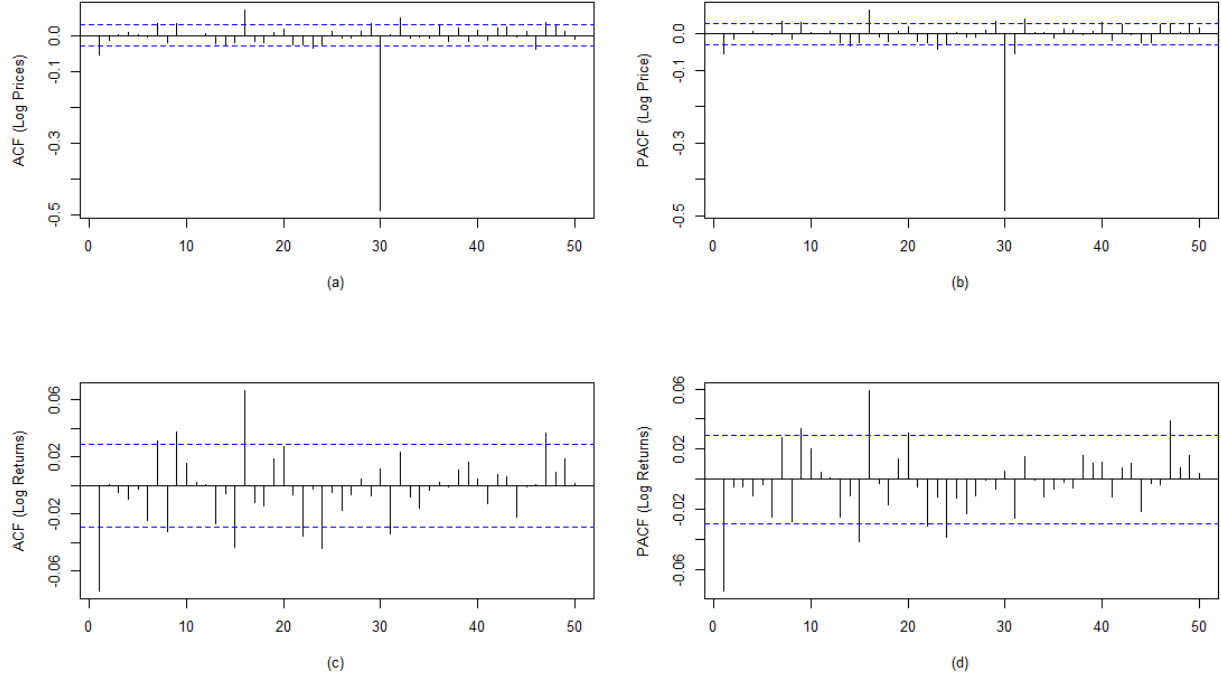


Figure 3.1: ACF and PACF for the log returns/prices of QCOM

past forecasting errors of the $MA(q)$ model. We define the model as follows:

$$r_t = \phi_0 + \sum_{i=1}^p \phi_i r_{t-i} + a_t - \sum_{i=1}^q \theta_i a_{t-i}, \quad (3.1)$$

where $\{a_t\}$ is a white noise series and p and q are nonnegative integers. Since the ACF and PACF are not informative in determining the order of an ARMA model, we use the extended autocorrelation function (EACF) to specify the order of an ARMA process. If we have a consistent estimate of the AR component of an ARMA model, we can derive the MA component. From the derived MA series, we can use ACF to identify the order of an MA component.

Table 3.1: Extended Autocorrelation Function (EACF) for QCOM Log Returns

(p, q)	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	o	o	o	o	x	x	x	x	o	o	o	o	o
1	x	o	o	o	o	o	o	o	o	o	o	o	o	o
2	x	o	o	o	o	o	o	o	o	o	o	o	o	o
3	x	o	o	o	o	o	o	o	o	o	o	o	o	o
4	x	x	o	o	o	o	o	o	o	o	o	o	o	o
5	x	x	x	o	o	o	o	o	o	o	o	o	o	o
6	x	x	x	x	o	o	o	o	o	o	o	o	o	o
7	x	x	x	x	x	o	o	o	o	o	o	o	o	o

Unlike the standard ACF and PACF – which are more effective for pure AR or MA models – the EACF is especially useful when dealing with mixed ARMA models. The output of an EACF is a two-way table, where the rows correspond to AR order p and the columns to MA order q . Table (3.1) shows the Theoretical EACF for an ARMA model, where X denotes statistically significant autocorrelation, and O denotes statistically insignificant autocorrelation, which suggests a good candidate for the true model. In general, for an ARMA(p, q) model, the triangle of O will have its upper left vertex at the (p, q) position, which we can see is an ARMA(1,1) model. This will be the basis for our model selection.

3.1 Model Selection

ARMA(1,1) is the basis for our models, but as we can see from the EACF, we can also consider entries that fall inside the "O triangle" (starting at $p = 1, q = 1$). The residuals at these points are also insignificant and potentially valid model candidates. Considering this, we also consider ARMA(1, 2), ARMA(2, 1), ARMA(2, 2), ARMA(1, 3), and ARMA(0, 2). We have picked a combination of models that either add an extra AR or MA term to balance our selection criteria.

Table (3.2) the selected fit for both the log returns and the log prices. The log prices includes both seasoning and regular differencing. For the selection criteria we use the AIC, which indicates a better-fitting model with a balance between the goodness of fit and model complexity. We choose the model with the smallest AIC, which is the ARMA(2,2) for the log prices and ARMA(0,2) for the log returns.

Table 3.2: AIC for Log Prices and Log Returns

	ARMA(1,1)	ARMA(2,1)	ARMA(2,2)	ARMA(1,3)	ARMA(0,2)
Log Prices	-21945.69	-21943.79	-21943.32	-21941.81	-21945.81
Log Returns	-21958.18	-21956.2	-21954.39	-21954.37	-21958.15

3.2 Forecasting

Like the behavior of ACF, forecast of an ARMA(p, q) model have similar characteristics as those of an AR(P) model after adjusting for the impacts of the MA component. The form for the ℓ -step-ahead forecast is the following:

$$\hat{r}_h(\ell) = E(r_{h+\ell}|F_h) = \phi_0 + \sum_{i=1}^p \phi_i \hat{r}_h(\ell - i) - \sum_{i=1}^q \theta_i a_h(\ell - i) \quad (3.2)$$

Figure (3.2) presents a 1000-step ahead forecast for both the log prices (left panel) and log returns (right panel) of QCOM based on fitted ARIMA models.

The left panel shows the forecast from an ARIMA(2,1,2)(0,1,0) model applied to the log price series. The forecasted trajectory exhibits a flattening trend with widening confidence intervals over time, reflecting increasing uncertainty as the forecast horizon extends. Despite recent upward momentum in the log price, the model projects a relatively stationary mean path, suggesting that much of the recent trend may not persist in the long term. The fan-shaped prediction intervals highlight the stochastic nature of the process, primarily due to the seasonal and regular differencing applied.

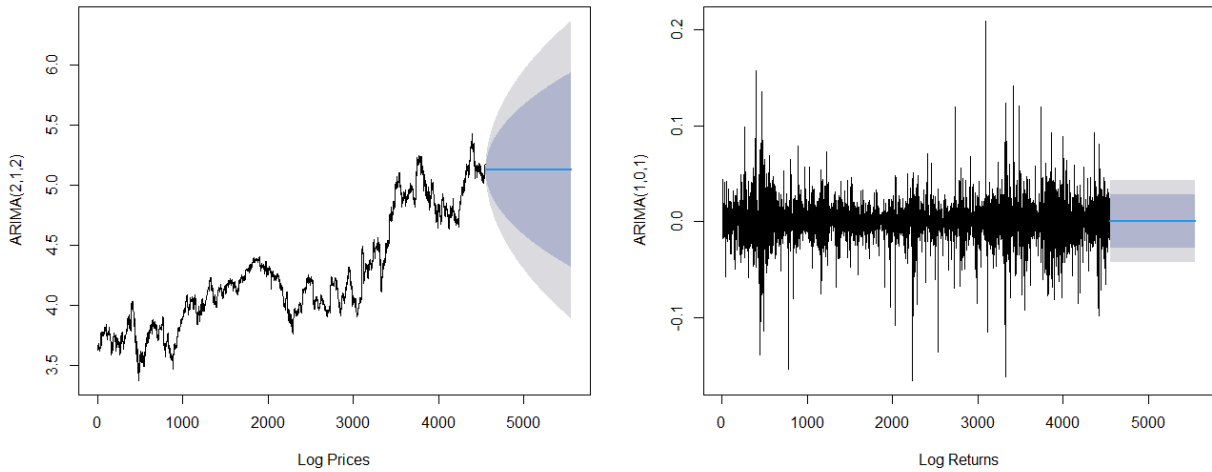


Figure 3.2: 1000 Step Ahead Forecast for the Log Returns and Prices

Figure (3.3) shows the residual diagnostics for the fitted ARIMA(2,1,2) model applied to the QCOM log returns. The right panel displays the forecast from an ARIMA(0,1,1) model fitted to the log returns. The forecasted returns converge quickly to a mean close to zero, with narrow and stable confidence bands. This behavior aligns with the efficient market hypothesis, where log returns are assumed to be white noise, and thus, future returns are unpredictable based on past behavior. The low variance and flat mean reinforce the conclusion that the return process is stationary and lacks significant autocorrelation beyond short lags.

The right panel displays the forecast from an ARIMA(0,1,1) model fitted to the log returns. The forecasted returns converge quickly to a mean close to zero, with narrow and stable confidence bands. This behavior aligns with the efficient market hypothesis, where log returns are assumed to be white noise, and thus, future returns are unpredictable based on past behavior. The low variance and flat mean reinforce the conclusion that the return process is stationary and lacks significant autocorrelation beyond short lags.

Overall, the forecast results suggest that while log prices show medium-term variability and are influenced by both trend and seasonal components, the log returns remain largely mean-reverting with no significant drift, supporting the appropriateness of ARIMA models in modeling these two distinct dynamics. The top panel shows the standardized residuals

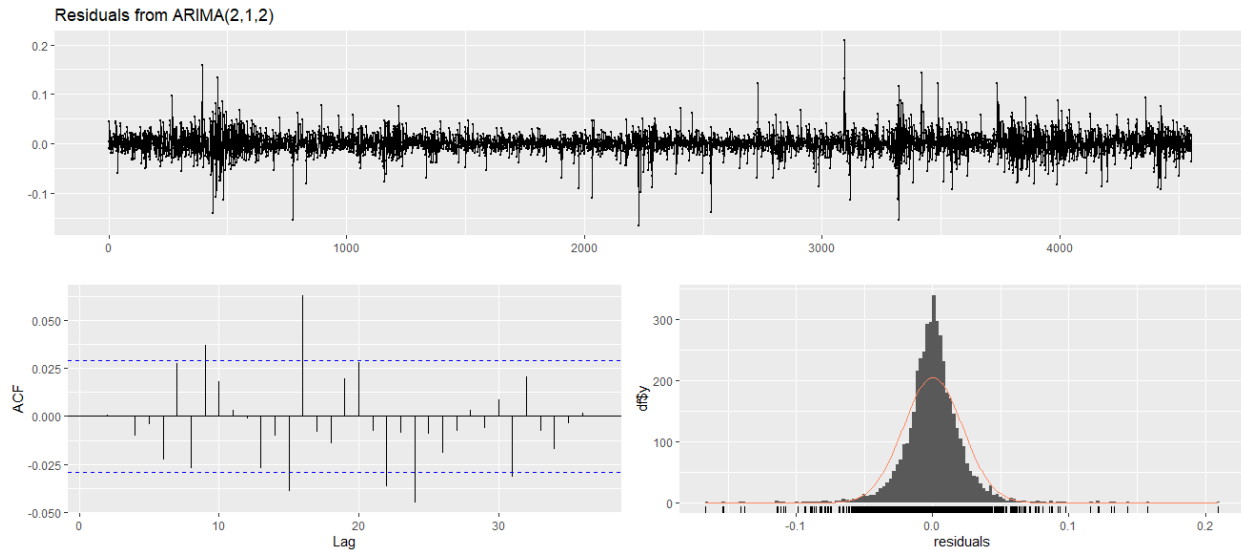


Figure 3.3: Residual Plots for the ARIMA(2,1,2) model for the Log Returns

over time. The residuals appear to fluctuate around zero with no visible trend or systematic pattern, suggesting that the model has effectively captured the underlying structure of the time series. Some sporadic large residuals are observed, but they are infrequent and do not persist.

The bottom-left panel shows the ACF of the residuals, where most autocorrelations lie well within the 95% confidence boundary. This indicates the absence of significant autocorrelation in the residuals, confirming that the ARIMA model has adequately removed time dependence from the series.

The bottom-right panel shows a histogram of the residuals overlaid with a normal density curve. The residuals are approximately centered at zero and exhibit a roughly bell-shaped distribution, although there is some evidence of slight fat tails compared to a normal distribution.

Overall, the diagnostic plots support the adequacy of the ARIMA(2,1,2) model. The residuals behave like white noise — they are uncorrelated and approximately normally distributed — suggesting that the model is a good fit for the data and suitable for forecasting purposes. Figure (A.1) also exhibits the same residual distribution for the ARIMA(1,0,1) model with the log prices.

Chapter 4

Vector Autoregressive Models

4.1 VAR Analysis

A Vector Autoregressive model of order p , denoted as $\text{VAR}(p)$, is used to model multivariate time series where each variable depends linearly on its own past values and those of other variables in the system. In contrast to univariate autoregressive models, VAR models account for the interdependencies among multiple time series by allowing each variable in the system to depend not only on its own past values but also on the past values of all other variables. This makes VAR particularly suitable for analyzing financial and economic systems where variables are often jointly determined, such as stock returns.

For k -dimensional time series r_t , the $\text{VAR}(p)$ model is specified as

$$r_t = \phi_0 + \Phi_1 r_{t-1} + \cdots + \Phi_p r_{t-p} + a_t \quad (4.1)$$

where Φ_0 is a $k \times 1$ vector of constants, Φ_j is a $k \times k$ autoregressive coefficient matrix, and a_t is a white noise innovation vector $\mathbb{E}[a_t] = 0$ and $\text{Cov}(a_t) = \Sigma_a$, a positive definite matrix.

One of the key strengths of VAR models is their ability to facilitate impulse response analysis, which tracks how a shock to one variable affects the entire system over time. This is useful in understanding transmission mechanisms, such as how a central bank interest rate change impacts inflation and output. Another common application is forecast error variance decomposition, which quantifies the contribution of each variable's shock to the forecast error variance of others, offering insights into relative importance.

Figure (4.1) shows a panel of scatterplots, which illustrates the joint behavior and possible dynamic interactions between the monthly log returns of AMD and QCOM. In plot (a), which compares the concurrent log returns of QCOM and AMD, we observe a moderate positive association. This indicates that when AMD experiences a positive or negative return in a

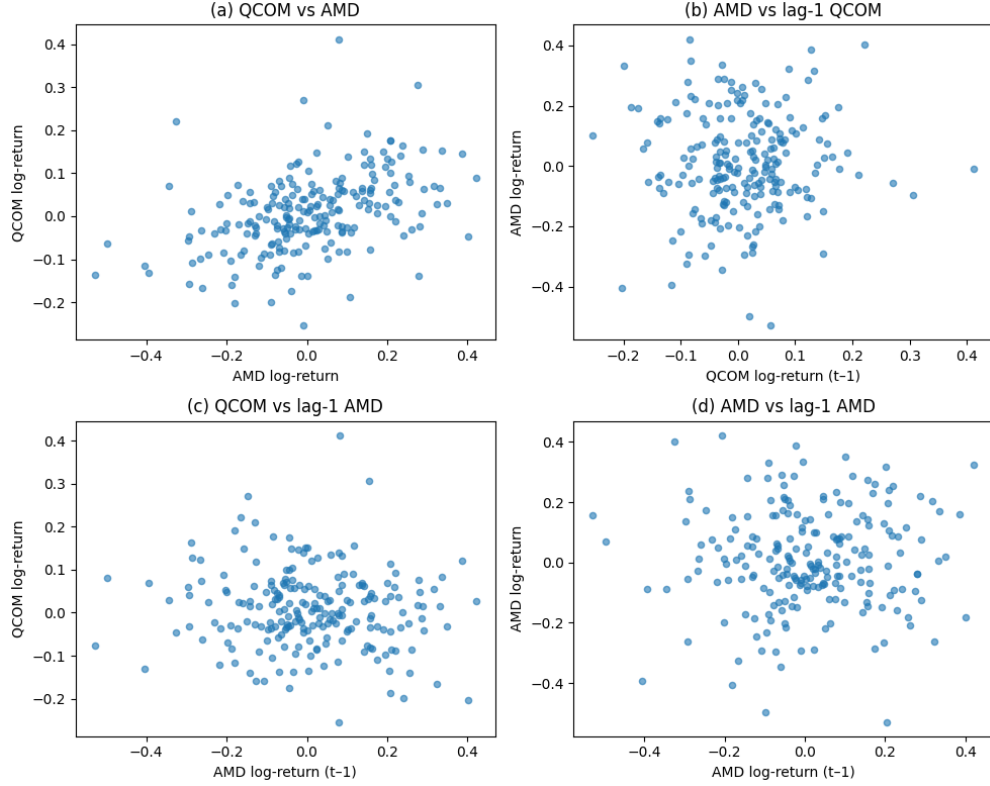


Figure 4.1: Scatter Plot of Monthly log returns of AMD and QCOM

given month, QCOM tends to move in the same direction. This contemporaneous correlation suggests shared exposure to common market factors or industry-specific dynamics (both being semiconductor firms).

Plot (b) investigates whether QCOM's past return (lag-1) helps explain AMD's current return. The scatter shows weak structure and appears diffuse, suggesting limited predictive power from QCOM's previous month return on AMD's current return. Likewise, plot (c) compares QCOM's current return to AMD's lag-1 return, and again, no strong relationship emerges, indicating that AMD's past performance does not substantially inform QCOM's present movement.

Finally, plot (d) looks at AMD's return against its own lag-1 value. The spread suggests minimal autocorrelation, which aligns with the empirical observation that individual stock returns are typically weakly autocorrelated at monthly frequencies due to market efficiency.

In our research, we estimate a VAR(1) model using monthly log returns from QCOM (Qualcomm) and AMD (Advanced Micro Devices). The VAR(1) specification implies that each variable in the system is modeled as a linear function of one lag of both itself and the other variable, plus a constant and a random error term. This structure allows us to explore how past performance in one stock may influence future returns in both, offering insights

into potential lead-lag relationships and co-movement dynamics in financial markets.

Matheamatically, a VAR(1) model for two variables can be expressed as the following:

$$r_t = \phi_0 + \Phi_{r_{t-1}} + a_t \quad (4.2)$$

where ϕ_0 is a k -dimensional vector, Φ is a $k \times k$ matrix, and $\{a_t\}$ is a sequence of serially uncorrelated random vectors with mean zero and covariance matrix Σ . The VAR(1) model consist of the following two equations:

$$\begin{aligned} r_{1t} &= \phi_{10} + \Phi_{11}r_{1,t-1} + \Phi_{12}r_{2,t-1} + a_{1t}, \\ r_{2t} &= \phi_{20} + \Phi_{21}r_{1,t-1} + \Phi_{22}r_{2,t-1} + a_{2t}, \end{aligned} \quad (4.3)$$

The individual equations derived from the matrix form are QCOM returns:

$$\begin{aligned} \text{QCOM}_t &= 0.002356 \cdot \text{QCOM}_{t-1} + 0.03490 \cdot \text{AMD}_{t-1} + 2.1352 \\ \text{AMD}_t &= -0.03611 \cdot \text{QCOM}_{t-1} - 0.08491 \cdot \text{AMD}_{t-1} + 1.9618 \end{aligned} \quad (4.4)$$

which gives us the following matrix representation

$$\begin{bmatrix} \text{QCOM}_t \\ \text{AMD}_t \end{bmatrix} = \begin{bmatrix} 2.1352 \\ 1.9618 \end{bmatrix} + \begin{bmatrix} 0.02356 & 0.03490 \\ -0.03611 & -0.08491 \end{bmatrix} \begin{bmatrix} \text{QCOM}_{t-1} \\ \text{AMD}_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{\text{QCOM},t} \\ \epsilon_{\text{AMD},t} \end{bmatrix} \quad (4.5)$$

The QCOM equation shows weak positive autocorrelation and a mild positive effect from AMD's previous return, suggesting slight predictive influence. The AMD equation exhibits negative coefficients for both lagged variables, indicating possible mean-reversion and a modest inverse relationship with QCOM's past performance. The intercept terms represent average return levels adjusted for lagged influences.

Overall, the VAR(1) model captures the dynamic feedback between AMD and QCOM returns, allowing for richer modeling than univariate time series approaches. The small magnitude of the coefficients suggests that while interdependence exists, the direct predictive power is limited. However, even weak relationships can be valuable for short-term forecasting, constructing hedging strategies, or understanding market behavior under shocks (e.g., via impulse response functions).

4.2 Cointegration

In financial time series analysis, it is common to observe that individual asset prices or economic indicators follow nonstationary behavior, typically modeled as integrated processes

of order one, $I(1)$. However, economic theory often suggests that certain groups of such variables move together in the long run, maintaining stable equilibrium relationships despite short-term fluctuations. This phenomenon is known as *cointegration*. Formally, if x_{1t} and x_{2t} are both $I(1)$ series, they are cointegrated if there exists a scalar γ such that the linear combination $w_t = x_{1t} - \gamma x_{2t}$ is stationary ($I(0)$). The vector $(1, -\gamma)$ is then called the cointegrating vector, and w_t represents the equilibrium error or spread.

To formally test and model cointegrated systems, particularly in higher dimensions, the *Vector Error Correction Model (VECM)* framework is employed. This begins by rewriting a $\text{VAR}(p)$ model of nonstationary series X_t in terms of first differences, yielding the VECM:

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{p-1} \Phi_i^* \Delta X_{t-i} + a_t \quad (4.6)$$

Here, the matrix $\Pi = \Phi_1 + \dots + \Phi_p - I$ captures the long-run relationships, while Φ_i^* represent short-run dynamics. The key insight is that the rank of Π determines the cointegration structure: if $\text{rank}(\Pi) = r < k$, then there exist r cointegrating vectors, and we can factor Π as $\alpha\beta'$, where β contains the cointegrating vectors and α reflects the speed of adjustment toward long-run equilibrium.

Cointegration has significant implications in financial applications. For instance, in pairs trading, if two stock prices are cointegrated, the spread $w_t = p_{1t} - \gamma p_{2t}$ will be mean-reverting, allowing traders to exploit deviations from the equilibrium for profit. A trading strategy might involve taking a long position in one stock and a short position in the other when the spread deviates from its historical mean by a certain threshold.

The Johansen test, based on estimating the rank of Π in the VECM, is the standard method for identifying such cointegrating relationships in multivariate systems. The hypothesis test is as follows:

$$\begin{aligned} H_0 : \text{rank}(\Pi) &\leq r \\ H_a : \text{rank}(\Pi) &> r \end{aligned} \quad (4.7)$$

So increasing the assumed rank and testing how many stationary combinations exists.

Figure (A.2) shows the cumulative returns for both QCOM and AMD. The two series seem to track each other, especially in the later half of the sample. This suggest that a stationary spread may exists ($\omega_t = \text{QCOM}_t - \gamma \cdot \text{AMD}_t$)

Chapter 5

ARCH/GARCH Volatility Models

5.1 ARCH

An ARCH (autoregressive conditionally heteroscedastic) model is a model for variance of time series. ARCH models are used to describe a changing, possibly volatility variance. models were created in the context of econometric and finance problems having to do with the amount that investments or stock increases (or decrease) per time period, so there's a tendency to describe them as models for that type of variable. For that reason, the variable of interests will be log returns.

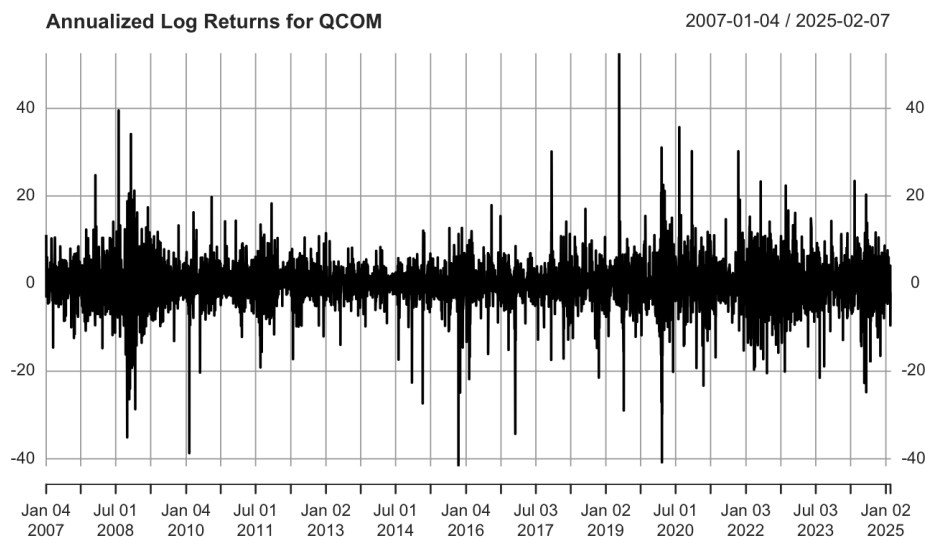


Figure 5.1: Annualized Logged Returns of QCOM

The ARCH(1) model for the variance of model r_t is that conditional on r_{t-1} , the variance at time t is

$$\text{Var}(r_t | r_{t-1}) = \sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 \quad (5.1)$$

Figure (5.1) shows the annualized log returns for QCOM from January 4, 2007, to February 7, 2025. The returns are highly volatile, with pronounced spikes-both positive and negative-especially noticeable during periods like the 2008 financial crisis and the COVID-19 pandemic in 2020. Overall, the data oscillates around zero, reflecting frequent fluctuations in QCOM's stock returns over the observed period. An ARCH model could be used for any series that has periods of increased or decreased variance. This might, for example, be a property of residuals after an ARMA or ARIMA model has been fit to the data. As such, we will use the ARMA(2,2) that we have built earlier.

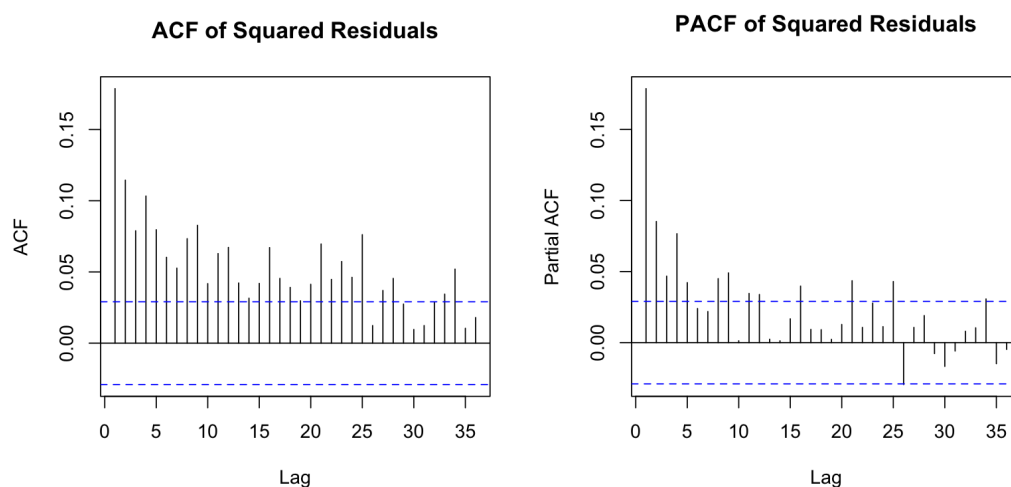


Figure 5.2: ACF and PACF of the Squared Residuals of ARMA(0,2)

Figure (5.2) displays the ACF/PACF in the ARMA(2,2) model residuals. For the ACF plot, we see multiple significant autocorrelations that extend well above the significance bounds. The first several lags (particularly lags 1-5) show strong positive autocorrelation. There is also persistence of significant spikes throughout many lags that indicate strong volatility clustering. There is also a decaying pattern, which is typical of financial time series with conditional heteroskedasticity. The PACF plot of the right shows strong significant spikes at early lags, particularly from lags 1-3. There are also several additional significant lags scattered throughout the plot.

The plot provides clear evidence of the ARCH effects in QCOM. this is confirmed by the The ARMA(0,2) model alone is not enough to capture the time-varying volatility in the QCOM time series, as there are significant autocorrelations extending across multiple lags. This suggest that higher-order ARCH model would be needed. Alternatively, we could use a GARCH-type model, which would provide more parsimonious fit given the persistent nature of the autocorrelation.

5.2 GARCH

An ARCH(m) proces is one for which the variance at time t is conditional on observations at the previous m times, and the relationship is

$$\text{Var} (r_t | r_{t-1}, \dots, r_{t-m}) = \sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \dots + \alpha_m r_{t-m}^2 \quad (5.2)$$

With certain constraints imposed on the coefficients, the r_t series squared will theoretically be $AR(m)$. On the other hand, GARCH (generalized autoregressive conditionally heteroscedastic) model uses values of past squared observations and past variances to model the variances at time t .

The GARCH model is an extension of the ARCH(p) model. We notice the additional term shown above when defining the conditional variance (volatility σ_t^2 at time t), which allows for modeling the conditional variance to be dependent on lagged versions of itself, using a lienar combination of $\alpha_0, \alpha_1, \dots, \alpha_n$ and the conditional variances at previous times. This addition ot the model statement makes GARCH models more flexible and able to capture the persistance of volatility. We first start by building the GARCH(1,1) model using the mean of the ARMA(0,2) model.

5.2.1 GARCH(1,1)

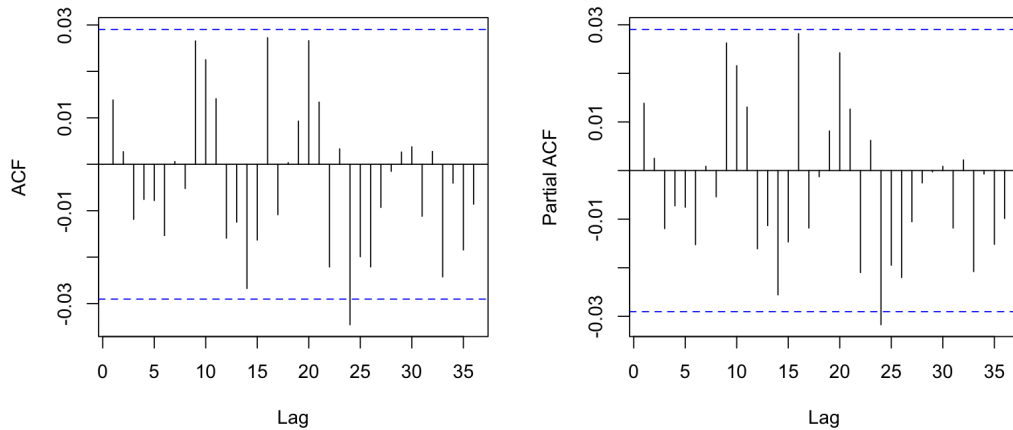


Figure 5.3: ACF and PACF of the Standard Residuals of GARCH(1,1) with ARMA(0,2) mean

Figure (5.3) shows the ACF/PACF standard residuals from GARCH(1,1) with ARMA(0,2) mean. The ACF plot shows that most autocorrelations are within the significance bounds,

with only a few minor spikes (lag 5). Otherwise, there are no systematic patterns of significant autocorrelation across lags. The PACF again shows a significant spike at lag 5, but all other lags fall within the confidence bounds. The lack of significant autocorrelation in both plots suggest that the ARMA(0,2) specification has adequately captured the linear dependence in the means of QCOM. The isolated spikes at lag 5 are likely due to random noise rather than model misspecification.

Figure (5.4) shows the ACF/PACF plots for squared residuals of the GARCH(1,1) model. A notable significant spike at lag 1 that clearly exceeds the confidence bounds. We also have some other minor spikes at various lags that approach but mostly stay within the significance boundaries. In terms of the PACF, we have multiple significant lags at 3.

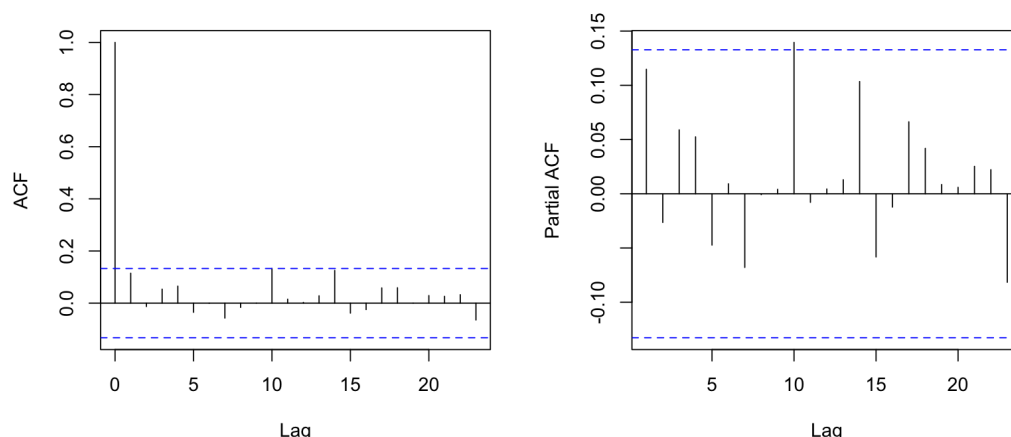


Figure 5.4: ACF and PACF of the Squared Residuals of GARCH(1,1) with ARMA(0,2) mean

This shows that there may be some remaining ARCH effects in the model, and more refinement is necessary.

We find the optimal GARCH orders by using the EACF, which involves locating the corner of insignificant autocorrelations (marked o); a similar process for identifying the optimal ARIMA model. As we can see Table (5.1) shows the first clear triangle which starts at AR=1 (GARCH Term) and MA=4 (ARCH term), suggesting a GARCH(1,1) specification, which is our first model of choice.

Table 5.1: Extended Autocorrelation Function (EACF) for QCOM Log Returns

(p, q)	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	o	o	o	o	o	o	o	o	o	o	o	o	o	o
1	x	o	o	o	o	o	o	o	o	o	o	o	o	o
2	x	x	o	o	o	o	o	o	o	o	o	o	o	o
3	x	x	x	o	o	o	o	o	o	x	o	o	o	o
4	x	x	x	x	o	o	o	o	o	o	o	o	o	o
5	x	x	x	x	o	o	o	o	o	o	o	o	o	o
6	x	x	x	o	x	o	o	o	o	o	o	o	o	o
7	o	o	x	o	x	o	o	o	o	o	o	o	o	o

5.3 Forecasting

An ARMA(2,2)-GARCH(1,1) model combines an Autoregressive Moving Average (ARMA) model for the conditional mean of a time series with a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model for its conditional variance. The mathematical derivation for the ARMA(2,2)

$$\begin{aligned}
 r_t &= \mu + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \epsilon_t \\
 &= 1.74405 - 0.05736 r_{t-1} + 0.87125 r_{t-2} - 0.11683 \epsilon_{t-1} - 0.98070 \epsilon_{t-2} + \epsilon_t
 \end{aligned} \tag{5.3}$$

where r_t is the value of the time series at time t , μ is the constant term, ϕ_1 and ϕ_2 are the autoregressive coefficients for the first and second lags of r_t , θ_1 and θ_2 are the moving average coefficients for the first and second lags of the error term ϵ_t

$$\begin{aligned}
 \sigma_t^2 &= \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \\
 &= 31.9997 + 0.04625 \epsilon_{t-1}^2 + 0.89064 \sigma_{t-1}^2
 \end{aligned} \tag{5.4}$$

where σ_t^2 is the conditional variance at time t , ω is the constant term in the variance equation, α_1 is the ARCH parameter, which measures the reaction to conditional variance to past shocks, and β_1 is the GARCH parameter, which measures the persistence of the conditional variance based on past conditional variances.

The term ω term (32.000), representing a baseline variance, is not statistically significant ($p=0.118$). The ‘alpha1’ coefficient (0.046), which captures the impact of past shocks on volatility, is also not statistically significant ($p=0.137$), suggesting a weak immediate reaction

of volatility to new information. However, the ‘beta1’ coefficient (0.891) is highly significant ($p=0.000$) and close to 1, indicating strong persistence in volatility; past volatility levels are a key determinant of current volatility, causing volatility clustering. The sum of α_1 and β_1 ($0.046 + 0.891 = 0.937$) is high, confirming that shocks to volatility are quite persistent, though the sum being less than 1 implies the variance process is stationary.

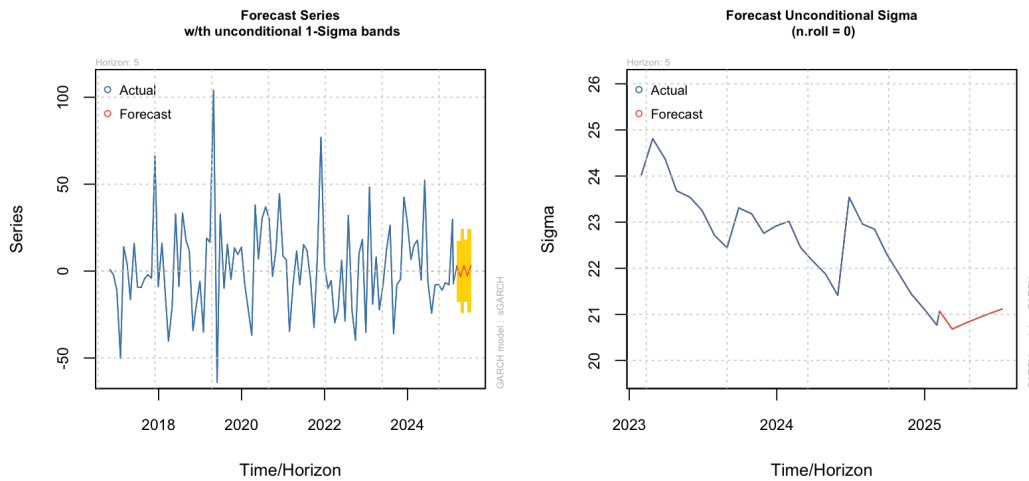


Figure 5.5: ACF and PACF of the Squared Residuals of GARCH(1,1) with ARMA(0,2) mean

Figure (5.5) shows the forecast for the GARCH(1,1) model with ARMA(2,2) mean, which provides valuable insights into both the predicted returns and volatility patterns. The left panel displays the historical series (blue line) characterized by significant fluctuations over the 2018-2025 period, with several dramatic spikes exceeding 50 units in both positive and negative directions. The forward-looking forecast (orange line) projects relatively modest returns with a slight upward trend, accompanied by widening yellow confidence bands that reflect increasing uncertainty over the 5-period forecast horizon. This suggests that while the model anticipates positive returns in the near future, it acknowledges considerable uncertainty in these predictions, consistent with the historical volatility observed in the series.

The right panel, displaying unconditional sigma (volatility) forecasts, reveals that volatility has followed a generally declining trend from about 24.5 units in early 2023 to approximately 20.7 units by early 2025. The model forecasts a modest uptick in volatility over the next five periods. This slight increase in projected volatility aligns with the expanding confidence intervals seen in the returns forecast. The overall pattern suggests that while the GARCH model expects volatility to remain lower than historical peaks, it anticipates some volatility persistence - a typical characteristic of financial time series modeled with GARCH specifications. This measured increase in forecasted volatility indicates the model’s response

to recent market conditions, though the magnitude remains relatively contained compared to historical levels.

Appendix A

Figures

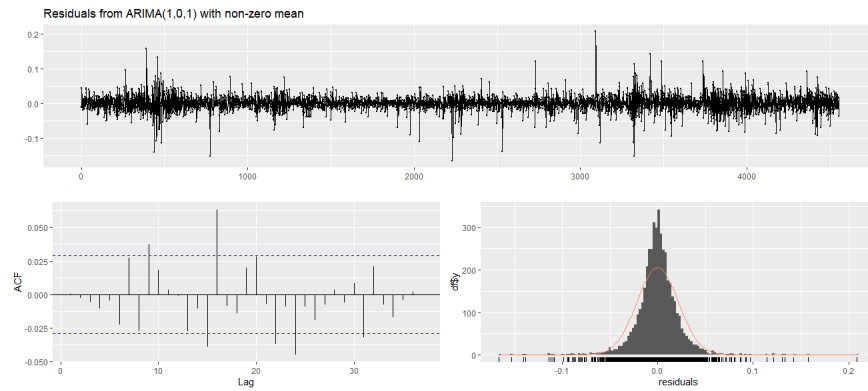


Figure A.1: Residual Plots for the ARIMA(1,0,1) model for the Log Prices

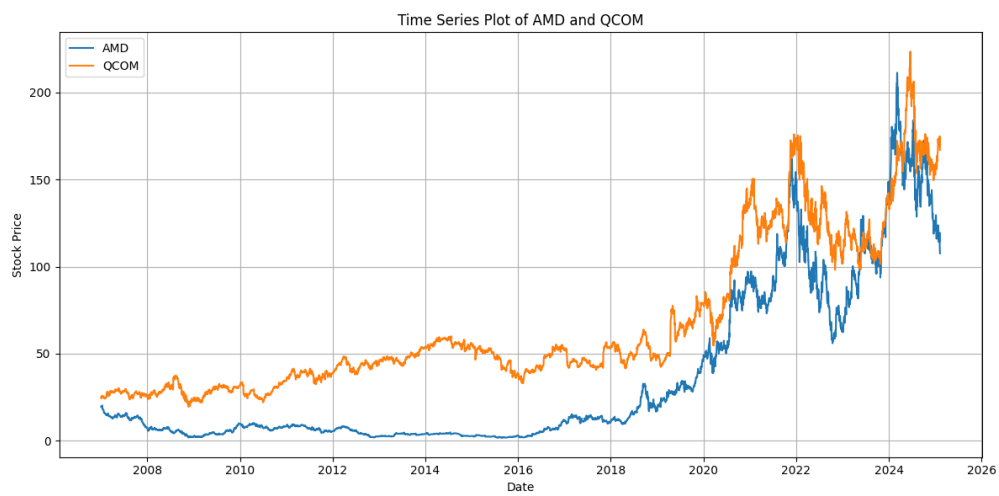


Figure A.2: Cumulative Returns of QCOM and AMD

Appendix B

Tables

Table B.1: Ljung-Box Test Model Fit

	Method	Lag order	t-statistic	p -value	Outcome
ARCH	Standard Residuals	10	4.5922	0.9167	No autocorrelation
GARCH	Squared Residuals	10	9.6976	0.4674	Stationary