Assignment 1 - Group 05

Duc Minh Nguyen, Tudor Ungureanu

March 28, 2025

Part 1

The aim of this research, is to analyze daily log-returns of Intel INTC stock, with a focus on autocorrelations, volatility clustering, and distributional assumptions.

Data

The time series considered in this paper is the daily log-return of the Intel stock (INTC), from 1992-02-25 to 2025-02-21. Denote it by $r_t = \Delta \log \mathrm{INTC}_t$, and its plot is given in Figure 1.

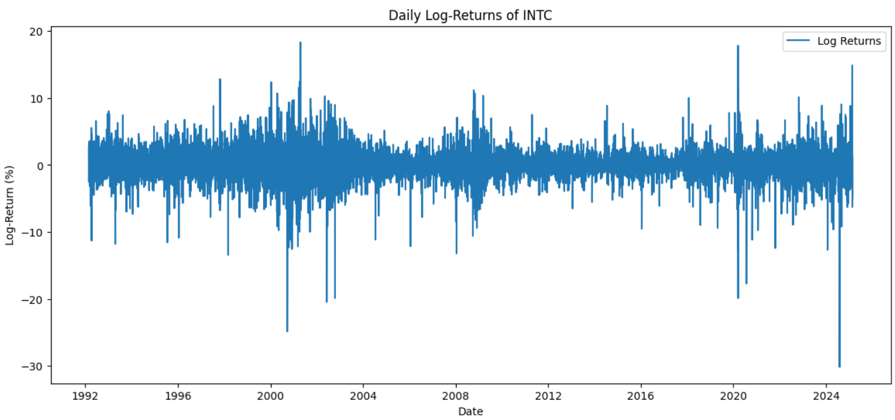

Figure 1: Daily log-return of INTC.

From the plot it is possible to observe some volatility clustering, notably in 2000-2004, 2008-2009, and 2019-2025 periods.

Model Selection

It is evident now that the model of r_t should account for volatility clustering, hence the GARCH model is selected to model the conditional variance. To further investigate, we have plotted the autocorrelation and partial autocorrelation functions of the log-return and squared log-return. Both the ACF and PACF of r_t cut-off at lag zero, and the rest of the lags are close to zero. This implies that the ARMA model should not suitable here, since for an ARMA model to

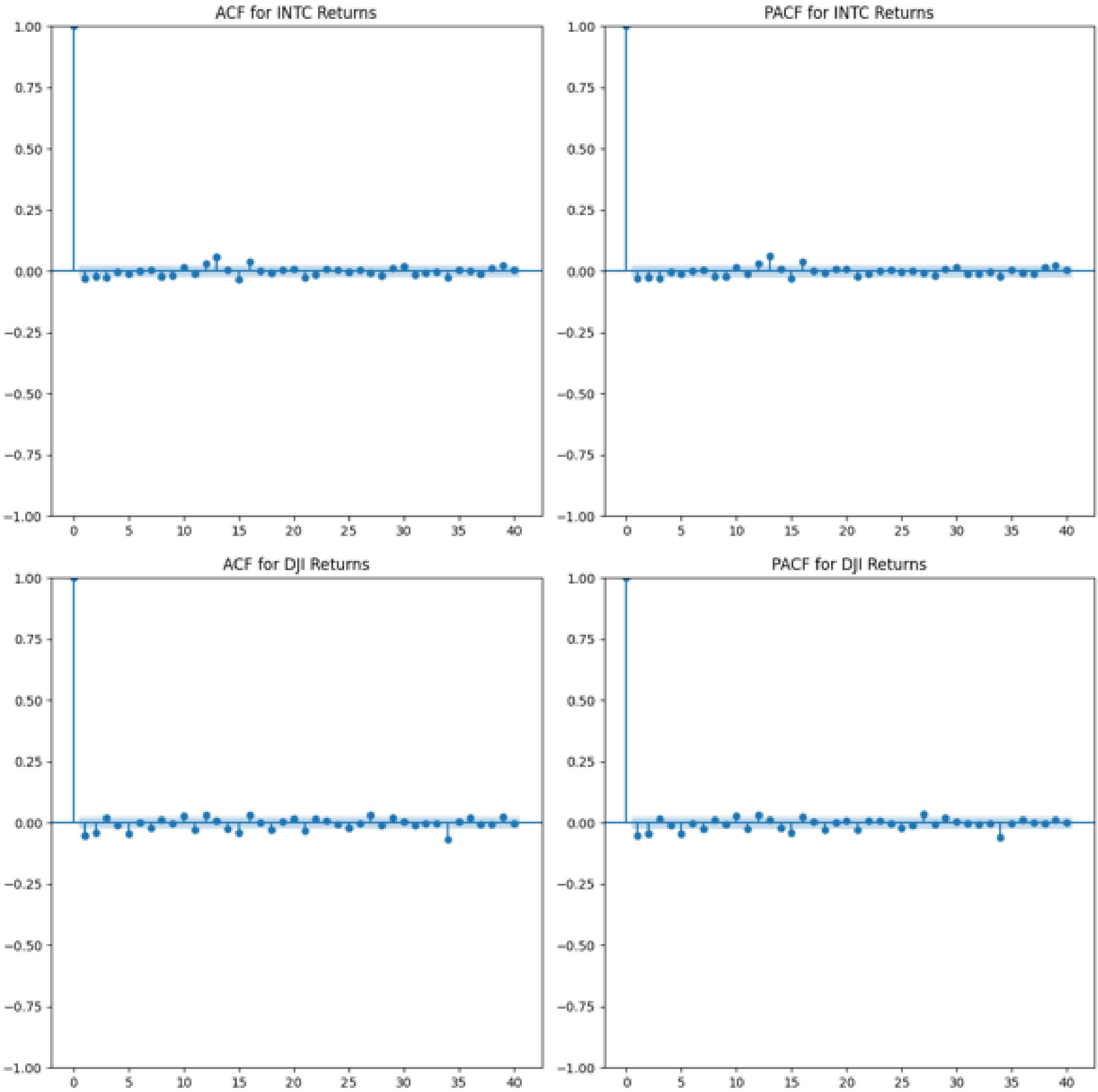

Figure 2: ACFs and PACFs.

make sence, both the ACF and PACF should decay exponentially. Instead, we model the mean process with a low order AR(p) model. Next, when looking at the log-return squared, it is possible to observe an exponential decay in both ACF and PACF. This suggests the usage of GARCH to model the conditional variance of r_t. Thus the model we select for the log-return of INTC stock is the AR(p)-GARCH(m, k):

$$r_t = \mu_t + a_t, \quad a_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \sim (0,1) \text{ white noise} \tag{1}$$

$$\mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i \mu_{t-i} + u_t, \quad u_t \sim \text{white noise} \tag{2}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{k} \beta_j \sigma_{t-j}^2 \tag{3}$$

The orders p, m, k of the model is selected based on Bayesian information criterion (BIC). The BIC for each $p \in \{1, ..., 5\}$ and $(m, k) \in \{(1,0), (1,1), (2,0), (2,1), (1,2), (2,2)\}$ are computed

Table 1: BIC values for different AR and GARCH orders

AR Order	GARCH Order	BIC
1	(1, 2)	36415.185489
2	(1, 2)	36418.874973
3	(1, 2)	36422.692213
1	(2, 2)	36424.210343
4	(1, 2)	36425.602779
1	(1, 1)	36427.526569
2	(2, 2)	36427.899707
5	(1, 2)	36428.572300
2	(1, 1)	36431.468233
3	(2, 2)	36431.716829
4	(2, 2)	36434.627271
3	(1, 1)	36435.215248
1	(2, 1)	36436.551613
5	(2, 2)	36437.596672
4	(1, 1)	36437.990585
2	(2, 1)	36440.492965
5	(1, 1)	36441.026144
3	(2, 1)	36444.239861
4	(2, 1)	36447.015077
5	(2, 1)	36450.053096
1	(2, 0)	37216.520481
3	(2, 0)	37219.916828
2	(2, 0)	37220.792148
4	(2, 0)	37224.568057
5	(2, 0)	37227.382763
2	(1, 0)	37639.184716
1	(1, 0)	37639.963818
3	(1, 0)	37641.909827
4	(1, 0)	37645.814170
5	(1, 0)	37649.005772

and displayed in ascending order in Table 1.

Based on BIC, we choose the AR(1)-GARCH$(1, 2)$ model.

Next, we investigate the distribution of the disturbance ε_t by looking at the histogram of r_t (Figure 3). It is possible to see that the data fits relatively well to a normal distribution, so we will first assume that the errors are normal. After estimation of the model, we will return to this by testing on the normality of the fitted residuals.

Estimation and Evaluation

The AR(1)-GARCH$(1, 2)$ model for r_t is estimated with the MLE, which yields $\hat{\phi}_0 = 0.0547$ and $\hat{\phi}_1 = -0.0132$ for the AR part, and $\hat{\omega} = 0.0415$, $\hat{\alpha}_1 = 0.0548$, $\hat{\beta}_1 = 0.3225$, and $\hat{\beta}_2 = 0.6169$ for the GARCH part.

Next, three tests are performed on the residual series of the model. The first test is the Ljung-Box test for serial correlation, with the null hypothesis that the residuals are not autocorrelated. The second test is the ARCH LM test on the conditional heteroskedasticity in the variance process, with the null hypothesis that there is no ARCH effect in the squared residual series. The last test is the Jarque-Bera normality test of the residual series. The p-values of these test are presented in Table 2. The Ljung-Box test's p-value is smaller than the 0.05 threshold, indicating that the residual series is autocorrelated. This is not expected, since the GARCH model should account for the serial correlation. It means that for our chosen model, the serial correlation still persists.

The p-value of ARCH-LM test is extremely small, which surely rejects the null of no conditional heteroskedasticity. This is again not in line with the fact that the GARCH model

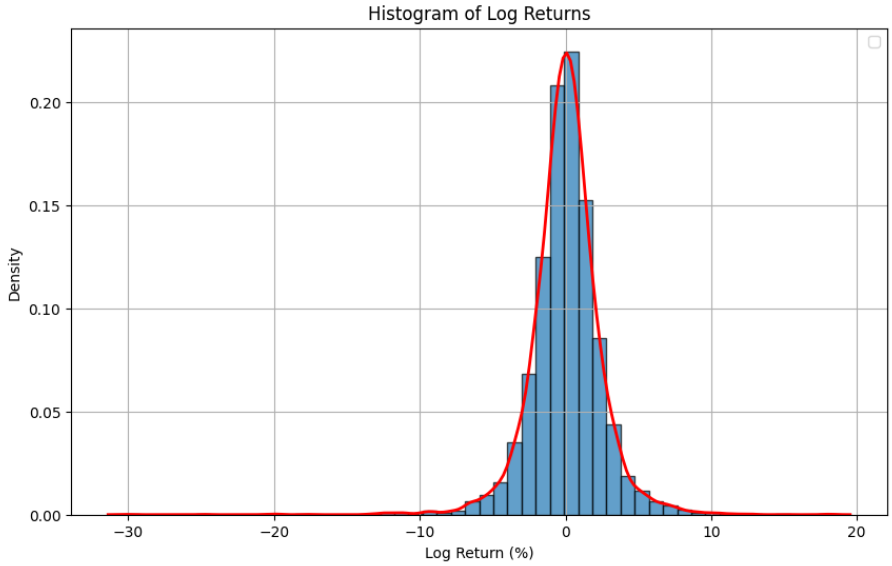

Figure 3: Histogram of log-return.

Test	p-value
Ljung-Box Test	0.013554
ARCH-LM Test	$6.7436\ 10^{-89}$
Jarque-Bera Test	0.0

Table 2: p-values of tests in the model with normal disturbance.

should clear the conditional heteroskedasticity effect in the squared residuals.

Finally, the JB test rejects the normality assumption of the residuals. This lead us to consider reestimating the log-return with another distributed errors. We want to account for the extreme events in the tail, since Intel is a tech company, which is sensitive to disruptions in the market. Thus, we assume a t-distributed error for the AR(1)-GARCH(1,2) model, and we reestimate it in the next section.

Reestimation

The same AR(1)-GARCH(1,2) model is considered, buth with $\varepsilon_t \sim t(\nu)$. This yields the estimates $\hat{\phi}_0 = 0.0748$ and $\hat{\phi}_1 = -8.2226 \cdot 10^{-3}$ for the AR part, and $\hat{\omega} = 0.0168$, $\hat{\alpha}_1 = 0.0504$, $\hat{\beta}_1 = 0.6147$, and $\hat{\beta}_2 = 0.3334$ for the GARCH part.

Test	p-value
Ljung-Box Test	0.005707
ARCH-LM Test	$3.18279 \cdot 10^{-89}$
Jarque-Bera Test	0.00

Table 3: p-values of tests in the model with $t(\nu)$ disturbance.

The tests results of reestimated model have not improved the serial correlation and conditional heteroskedasticity, as both LB and ARCH-LM p-values are still less than 0.05 threshold.

Conclusion

After estimating both the AR(1)-GARCH(1,2) with normal and t-distributed errors, the results still indicate a significant autocorrelation and ARCH effect present in the residuals. This indicate that the AR(1)-GARCH(1,2) model might be misspessified. For further research, we suggest to 1) use ARMA for the mean process, and 2) use models that accounts for leverage effects and other unobserved effects like E-GARCH or T-GARCH for the conditional variance.

Part 2

This part presents the backtesting results of the one-day-ahead 99% Value at Risk (VaR) forecast using an ARMA-GARCH model and a historical simulation approach. The estimation period begins before 2000, and the evaluation period is extended up to 2025. The effectiveness of the VaR estimates is assessed using both visual and statistical backtesting techniques.

Methodology

The ARMA-GARCH model was used to estimate conditional volatility, and the VaR one day ahead was calculated assuming a normal distribution of residuals. The historical simulation method was also applied, where past log-returns were used to construct an empirical distribution of losses. The performance of both methods was evaluated using:

- Basel Traffic Light Test
- Kupiec's Proportion of Failures (POF) Test
- Christoffersen's Independence Test

Results

The following figure presents the comparison between actual log-returns and the estimated VaR values.

From Figure 4, several observations can be made:

- The ARMA-GARCH VaR (red line) is more responsive to market fluctuations compared to the Historical VaR (black dashed line), particularly during high-volatility periods.
- During major market downturns, such as the COVID-19 crash in 2020 and recent declines in 2024, the ARMA-GARCH VaR adjusts significantly, capturing extreme events more effectively than Historical VaR.
- Both models exhibit exceedances where actual log-returns (blue line) drop below the estimated VaR. However, the number of exceedances for both models suggests that they may be underestimating risk.

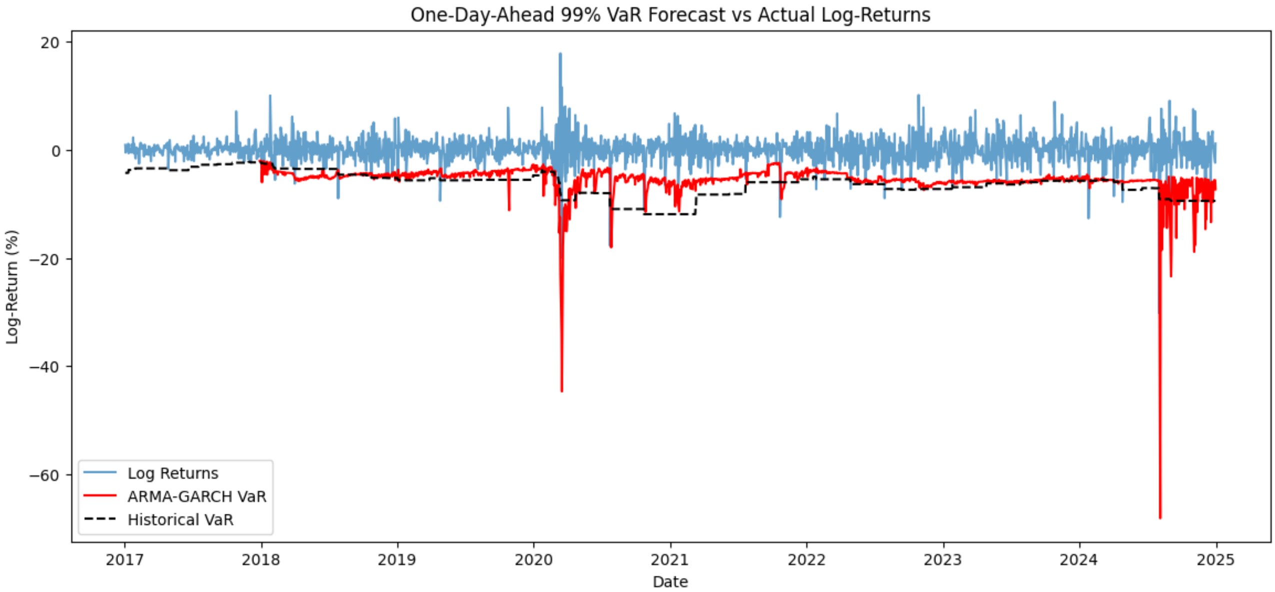

Figure 4: One-Day-Ahead 99% VaR Forecast vs. Actual Log-Returns

Backtest Results

Test	VaR Historical	VaR GARCH
Basel Traffic Light Test	Red Zone	Red Zone
Kupiec POF Test (Exceedances)	25	40
Kupiec POF Test (p-value)	0.0000	0.0000
Christoffersen Independence Test (LR)	260.6663	392.5997
Christoffersen Independence Test (p-value)	0.0000	0.0000

Table 4: Backtesting Results for VaR Models

Interpretation of Results

- **Basel Traffic Light Test:** Both the historical and ARMA-GARCH VaR models fall into the **Red Zone**. This classification suggests that the models exhibit an excessive number of exceedances (instances where the actual loss is greater than the forecasted VaR). According to Basel II and III regulations, a model in the Red Zone is likely misspecified and may not be suitable for accurate risk management.

- **Kupiec Proportion of Failures (POF) Test:** The p-values for both models are 0.0000, indicating that the null hypothesis of an accurate exceedance rate (consistent with the expected 1% exceedance for a 99% VaR) is strongly rejected. This suggests that both models systematically underestimate risk, leading to an underestimation of the frequency of extreme negative returns. The ARMA-GARCH model, with 40 exceedances compared to 25 in the historical model, appears to perform worse in this regard.

- **Christoffersen Independence Test:** The likelihood ratio (LR) values for both models are extremely high (260.6663 for historical VaR and 392.5997 for ARMA-GARCH VaR), with corresponding p-values of 0.0000. This indicates strong rejection of the null hypothesis that exceedances occur independently over time. Instead, exceedances appear to be clustered, meaning that periods of extreme losses tend to follow one another rather than occurring randomly.

- **Comparison of Models:** While both models are misspecified, the ARMA-GARCH model exhibits more exceedances, indicating an even greater underestimation of risk compared to the historical simulation approach. This may be due to the model's inability to fully capture the extreme tails of the return distribution, particularly in periods of high volatility.

Conclusion

The backtesting results indicate that both VaR models fail regulatory tests and are likely misspecified. The ARMA-GARCH model, while more dynamic in capturing volatility changes, still significantly underestimates risk, particularly during extreme market downturns. The presence of clustered exceedances suggests that both models fail to account for time-dependent risk, making them unreliable for real-world capital requirement calculations.

However, the extreme historical values observed during this period have a substantial impact on the results. If the dataset contains prolonged volatility spikes or crisis periods, a model that does not sufficiently account for heavy-tailed distributions will struggle to produce accurate risk estimates. The high exceedance counts indicate that the models do not fully capture tail risk, particularly in turbulent market conditions.