Group Coursework Submission Form (PA)

Specialist Masters Programme

Please list all names of group members: (Surname, first name) 1. Sanjay Desai, Prathmesh 2. Sato, Nobuyuki 3. Veera, Rohen	**GROUP NUMBER:** 4
MSc in: Quantitative Finance	
Module Code: SMM 271	
Module Title: Econometrics of Financial Markets	
Lecturer: Professor Giovanni Urga	**Submission Date:** 26th of March 2025

Declaration:

By submitting this work, we declare that this work is entirely our own except those parts duly identified and referenced in my submission. It complies with any specified word limits and the requirements and regulations detailed in the coursework instructions and any other relevant programme and module documentation. In submitting this work we acknowledge that we have read and understood the regulations and code regarding academic misconduct, including that relating to plagiarism, as specified in the Programme Handbook. We also acknowledge that this work will be subject to a variety of checks for academic misconduct.

We acknowledge that work submitted late without a granted extension will be subject to penalties, as outlined in the Programme Handbook. Penalties will be applied for a maximum of five days lateness, after which a mark of zero will be awarded.

Marker's Comments (if not being marked on-line):

Deduction for Late Submission:

Final Mark:

For Students:

Once marked please refer to Moodle for your final coursework grade, including your Peer Assessment grade.

1 Order of integration of the variables

On the basis of visual inspection, all four series exhibit stable upward trends over time typically observed in deterministic trend, while frequency of mean-reverting behaviors are modest and persistence of their own past values are also observed in their trajectory. Another notable feature is that distinct upward trends of lp1q and lp2q are evident only after around 2010, implying possible structural breaks having occurred around that period.

Based on these findings, to examine whether the four series are unit-root process or not, we implemented the Augmented Dickey-Fuller (ADF) test of the following form:

$$\Delta Y_t = \alpha + \beta t + \theta Y_{t-1} + \sum_{i=1}^{p} \Delta Y_{t-i} + u_t \tag{1}$$

where the null hypothesis is $H_0 : \theta = 0$, against the alternative hypothesis $H_1 : \theta < 0$. The test statistic is:

$$t_{ADF} = \frac{\hat{\theta}}{SE(\hat{\theta})} \tag{2}$$

which conforms to ADF critical values reflecting the null of non-stationarity and the functional form given above. $SE(\theta)$ denotes the standard error of the estimated θ. Note that for simplifying the discussions, we did not take into account the possible structural breaks by adding control variables in this test.

Firstly, we estimated the model 1 with $p = 4$ as the series are quarterly frequency. Then, judging based on whether the lagged variables and the trend have a t-value that exceeds the significance level of 5%, we omitted unnecessary variables. We also confirmed that the obtained models except $lp3q$ are the most parsimonious in terms of Akaike Information Criteria (AIC) up to $p = 4$, and for $lp3q$ we chose $p = 1$ as it showed the lowest AIC though the t-value was insignificant at 5% level. A summary of the results is shown in Table 1.

Series	**Number of lag terms**	**Inclusion of trend term**	t_{ADF}	**ADF critical value (5%; 1%)**	**Unit root process?**
lp1q	1	Yes	-1.459	-3.457; -4.057	Yes
lp2q	3	Yes	-4.108**	-3.458; -4.059	Yes
lp3q	1	Yes	-1.191	-3.457; -4.057	Yes
lp4q	1	Yes	-2.818	-3.457; -4.057	Yes

Table 1: ADF test results

Based on the critical values above, we conclude that lp1q, lp3q, and lp4q contain a unit root. As for lp2q, t_{ADF} value is marginally lower than 1% significance level, suggesting we should reject the null hypothesis based on the standard procedure. However, while the ADF test assumes the error term is white noise, separately examined misspecification tests also indicated serial correlation at 1% significance level, along with heteroskedasticity at 5% level. Considering these additional results and the fact that t_{ADF} is only marginally lower than the critical value, we decide not to reject the null for lp2q as well and conclude that all four series have a unit root.

As a next step, we conducted another ADF tests on the first differences of each series to determine the order of integration. This time, the number of lag terms were decided on the basis of ACF up to four lags and the trend term was omitted. The results are presented in Table 2.

Series	**Number of lag terms**	t_{ADF}	**ADF critical value (5%; 1%)**	**Stationary**
Dlp1q	1	-6.521**	-2.892; -3.501	Yes
Dlp2q	4	-6.703**	-2.893; -3.503	Yes
Dlp3q	0	-7.682**	-2.892; -3.5	Yes
Dlp4q	1	-7.773**	-2.892; -3.501	Yes

Table 2: Unit root test

In all four cases, the first differences are stationary at 1% significance level. Thus, we conclude that all four series are $I(1)$.

2 Multivariate cointegration analysis

As we confirmed that all the four series are $I(1)$, letting y_t denote a 4 by 1 column vector of the four series, the stationary representation the four series is expressed as:

$$\Delta y_t = \mu + \Pi y_{t-1} + \sum_{i=1}^{p} \Gamma_i \Delta y_{t-i} + u_t \tag{3}$$

Where u_t is a white noise process. The corresponding unrestricted form is given by:

$$y_t = \alpha + \beta t + \sum_{i=1}^{p+1} \phi_i y_{t-i} + u_t \tag{4}$$

Because (3) is a stationary expression, if the original four series are all independent and stationary, the rank of Π has to be 4 for the equation (3) to hold. On the other hand, if there are $K \leq 4$ independent and non-stationary series included in y_t, the rank of Π has to be reduced to $4 - K$ for the (3) to be stationary. Therefore, the rank of correctly estimated Π can be understood as the number of cointegrated relationship.

In the ADF test of Q1, we saw the number of lagged differences necessary for obtaining stationary representation was 1 to 3 depending on the series. Hence, we first estimated the VAR model corresponding to (4) with each $p + 1 = 2$ and 4 with a constant and a trend. Then, dummy variables corresponding to the shocks exceeding 2.5 standard deviations were added to augment any missed dynamics for describing the error as white noise. Finally, based on Π assumed from the restricted model (3) corresponding to the estimated unrestricted model (4), the rank test of the following form was implemented:

$$\lambda_{trace} = T \sum_{i=r+1}^{4} ln(1 - \hat{\lambda}_i)$$

where each λ and r denotes the eigenvalue and the tested rank of Π, under the null hypothesis that the number of cointegrated relationship is less than or equal to r. The results are assessed by the asymptotic critical values on the basis of the functional form of (3). The misspecification test results, rank test results and the number of parameters are given in Table (3), (4), and (5).

Test	p=2 (without dummies)			p=2 (4 dummies)			p=4 (without dummies)			p=4 (2 dummies)		
	statistic	results	p-value	statistic	results	p-value	statistic	results	p-value	statistic	results	p-value
AR 1-5	$F(80, 239)$	1.112	0.268	$F(80, 223)$	1.086	0.316	$F(80, 207)$	1.421	0.025*	$F(80, 199)$	1.169	0.1921
Normality	$Chi2(8)$	11.582	0.171	$Chi2(8)$	8.175	0.417	$Chi2(8)$	13.18	0.106	$Chi2(8)$	5.971	0.651
Hetero	$F(72, 281)$	1.850	0.000**	$F(72, 265)$	1.492	0.013*	$F(136, 221)$	1.053	0.364	$F(136, 213)$	0.909	0.725
RESET23	$F(32, 267)$	1.881	0.004**	$F(32, 252)$	1.607	0.025*	$F(32, 237)$	1.855	0.005**	$F(32, 230)$	1.836	0.006**

Table 3: Misspecification test results

Rank	$p = 2$ (without dummies)		$p = 2$ (4 dummies)		$p = 4$ (without dummies)		$p = 4$ (2 dummies)	
	Trace statistic	p-value	Trace statistic	p-value	Trace statistic	p-value	Trace statistic	p-value
0	65.34	0.035*	126.27	0.000**	78.093	0.002**	118.65	0.000**
1	34.5	0.27	78.164	0.000**	44.841	0.030*	71.321	0.000**
2	18.996	0.287	36.433	0.001**	22.523	0.124	35.989	0.001**
3	7.5584	0.299	14.289	0.023*	10.361	0.113	12.573	0.048*

Table 4: Summary of rank tests

	$p = 2$ (without dummies)	$p = 2$ (4 dummies)	$p = 4$ (without dummies)	$p = 4$ (2 dummies)
No. of parameters	40	56	72	80
No. of observations	93	93	93	93

Table 5: Number of parameters

As seen from the misspecification test results in Table (3), while all model indicated some degree of functional form error, normality was not rejected in all models. There was also autocorrelation of the significance level 5% in $p = 2$ without dummies, heteroskedasticity of the significance level 5% in $p = 2$ with 4 dummies, and heteroskedasticity of the significance level 1% in $p = 2$ without dummies. In terms of the significance of the error and the parsimoniousness of the model, these results suggest that models $p = 4$ with 4 dummies and $p = 2$ with 4 dummies can be understood as the most reliable models.

Table (4) suggests that these two models consistently rejected the null hypothesis of the rank ≤ 2 at 1% significance level and rank ≤ 3 at 5% significance level. This result implies all four series are stationary at 5% significance level, contrary to our assumption of all four series are $I(1)$. The possible explanation to this results are as follows:

- RESET test results indicates there was possible misspecification in the functional form. This implies there might be cointegrations not covered in the four series or structual breaks not covered, or other relationship than the cointegration might exist.
- The asymptotic critical values did not serve as adequate criteria against the degree of freedom given the large number of parameters against the observations.

As a remedy for the second point, Richards (1995) suggests multiplying the asymptotic critical value by "degree-of-freedom correction term" given as $T/(T - (number\,of\,variables) \times (number\,of\,lags))$. In our instance, this will increase the critical values about 1.1 times in the case of $p = 2$, and 1.2 times for $p = 4$. In $p = 4$ with 2 dummies case, as the p-value is marginally below 0.05 in $H_0 : r \leq 3$, this adjusted criterion casts doubts on the significance of the original result. Meanwhile, for all cases in $H_0 : r = 0$, this adjustment has only a minor impact. Therefore, we conclude that the number of cointegrated relationships is between 1 and 3. As a final remark to the result, we should note that the existence of cointegration contradicts to the efficient market hypithesis, and has potential to assume unnatural movement among the cointegrated stocks as pointed by Richards (1995). Hence, the implication of this result requires further careful assessment to assume the true economic relationship among them.

3 GARCH and MGARCH models

This analysis explores volatility dynamics and interdependencies among NVIDIA (Dlp1), Amazon (Dlp2), Microsoft (Dlp3), and Procter & Gamble (Dlp4) using multivariate GARCH models. I used daily frequency data in logarithmic form, analyzing first differences from August 2000 to August 2024.

Descriptive Analysis of Return Series

The series display typical features of financial returns: near-zero means, significant excess kurtosis, and some skewness. Technology stocks show considerably higher volatility than the consumer staples stock (P&G).

Table 6: Descriptive Statistics of Daily Returns

Stock	Mean (%)	Std. Dev. (%)	Skewness	Kurtosis
NVIDIA	0.096	3.51	-0.44	16.38
Amazon	0.073	2.92	0.48	17.52
Microsoft	0.039	1.82	0.06	12.06
P&G	0.027	1.18	-0.12	12.08

Three MGARCH Models: Implementation and Results

RiskMetrics Approach Implemented with standard decay factor $\lambda = 0.94$. All mean constants are positive but only marginally significant. Log-likelihood: 64,121.45.

Dynamic Conditional Correlation (DCC) Model Shows significant persistence parameters ($\alpha = 0.0059$, $\beta = 0.9932$). The sum (0.9990) indicates extremely high persistence in conditional correlations. Unconditional correlations reveal Microsoft-Amazon with highest correlation (0.449), while P&G-NVIDIA show lowest (0.158). Log-likelihood: 65,968.54.

Scalar BEKK Model Shows high volatility persistence ($b_1 = 0.987$) and significant ARCH effect ($a_1 = 0.144$). Some remaining ARCH effects in P&G's squared standardized residuals suggest incomplete capture of dynamics. Log-likelihood: 65,855.71.

Table 7: Model Comparison

Model	Log-likelihood	AIC	SIC
RiskMetrics	64,121.45	-20.48	-20.47
DCC	65,968.54	-21.06	-21.04
Scalar BEKK	65,855.71	-21.03	-21.03

Model Comparison

The DCC model provides the best fit with the highest log-likelihood and lowest information criteria. All models capture significant volatility persistence, with shocks having long-lasting effects. Tech stocks consistently show higher volatility than P&G, reflecting their more speculative nature. The correlation structure shows moderate correlations between stocks, higher among tech stocks, suggesting common risk factors.

4 Testing for contagion

Contagion Analysis: Forbes and Rigobon (2002)

Theoretical Framework

Financial contagion is a complex phenomenon defined by Forbes and Rigobon (2002) as "a significant increase in cross markets linkage after a shock to one country or a group of countries". The key distinction between contagion and simple interdependence lies in the statistically significant enhancement of market connections during periods of economic stress.

Methodological Approach

The analysis employs the Forbes-Rigobon heteroskedasticity-consistent correlation adjustment, formally expressed as:

$$\rho = \frac{\rho^*}{\sqrt{1+\delta(1-(\rho^*)^2)}} \tag{5}$$

Where:

- ρ is the adjusted correlation coefficient
- ρ^* is the Pearson's conditional correlation coefficient
- $\delta = \frac{\sigma_h^2}{\sigma_l^2} - 1$
- σ_h represents volatility during the crisis period
- σ_l represents volatility during the stable period

Empirical Findings

Our analysis examined the potential contagion from Microsoft to three key stocks: NVIDIA, Amazon, and Procter & Gamble. The investigation revealed nuanced patterns of market interconnection:

Table 8: Forbes-Rigobon (2002) Contagion Analysis Results

Stock	**Pre-crisis Correlation**	**Crisis Correlation**	**FR Adjusted Correlation**	**Interpretation**
NVIDIA	0.4598	0.4936	0.3966 (Test stat: 0.9168)	No statistically significant evidence of contagion (below 3.84)
Amazon	0.4705	0.5656	0.4628 (Test stat: 0.0156)	No statistically significant evidence of contagion (below 3.84)
Procter & Gamble	0.1783	0.4057	0.3201 (Test stat: 3.8413)	Statistically significant evidence of contagion (exceeding 3.84)

Theoretical Implications

The findings challenge simplistic linear models of market interactions by demonstrating that contagion is not a uniform phenomenon. We propose a generalized contagion test statistic:

$$FR(i \to j) = \frac{\rho_h - \rho_l^*}{\sqrt{var(\rho_h - \rho_l^*)}} \tag{6}$$

Where:

- ρ_h denotes the correlation during the crisis period
- ρ_l^* is the Pearson's correlation during the stable period

Conclusion

The empirical evidence reveals a nuanced landscape of financial contagion. While market correlations generally increased during the 2008 financial crisis, only Procter & Gamble demonstrated statistically significant contagion from Microsoft. The research contributes to our understanding of how financial shocks propagate, challenging oversimplified interpretations of market interdependence.

Higher Moments Contagion: Frey, Martin and Tang (2010)

Theoretical Framework

Fry et al. (2010) extend the concept of financial contagion beyond traditional correlation analysis by introducing a sophisticated methodology that examines changes in higher-order conditional moments of asset returns. This approach provides a more nuanced understanding of market interdependence during financial crises.

Methodological Approach

The methodology introduces two key test statistics, CS1 and CS2, which evaluate coskewness through different return combinations:

$$CS1(i \to j; r_i^1, r_j^2) = \frac{\hat{\psi}_y(r_i^1, r_j^2) - \hat{\psi}_x(r_i^1, r_j^2)}{\sqrt{(4\hat{\nu}_{y|x_i}^2 + 2)/T_y + (4\hat{\rho}_x^2 + 2)/T_x}}^2 \tag{7}$$

$$CS2(i \to j; r_i^2, r_j^1) = \frac{\hat{\psi}_y(r_i^2, r_j^1) - \hat{\psi}_x(r_i^2, r_j^1)}{\sqrt{(4\hat{\nu}_{y|x_i}^2 + 2)/T_y + (4\hat{\rho}_x^2 + 2)/T_x}}^2 \tag{8}$$

Where:

- r_i^1 is the return in market i
- r_i^2 is the squared return in market i
- T_x and T_y are sample sizes in pre-crisis and crisis periods
- $\hat{\rho}_x$ and $\hat{\rho}_y$ are conditional correlations
- $\hat{\nu}_{y|x_i}$ is the Forbes-Rigobon adjusted sample correlation

The coskewness function takes the form:

$$\hat{\psi}_y(r_i^m, r_j^n) = \frac{1}{T_y} \sum_{t=1}^{T_y} \left(\frac{y_{i,t} - \hat{\mu}_y^i}{\hat{\sigma}_y^i} \right)^m \left(\frac{y_{j,t} - \hat{\mu}_y^j}{\hat{\sigma}_y^j} \right)^n \tag{9}$$

Empirical Findings

Our analysis of the higher moments test for Microsoft's relationship with three stocks revealed:

Table 9: Frey et al. (2010) Higher Moments Contagion Analysis Results

Stock	**CS1 Test Statistic**	**CS2 Test Statistic**	**Interpretation**
NVIDIA	38.6982	8.7959	Significant changes in return distribution characteristics
Amazon	0.6086	4.7075	Moderate alterations in higher-order moment relationships
Procter & Gamble	103.9062	119.6122	Dramatic shifts in joint return distribution

Conclusion

The higher moments analysis reveals a complex picture of market interdependence during the 2008 financial crisis. Procter & Gamble stands out with the most dramatic changes in its joint return distribution, highlighting the nuanced nature of financial contagion that extends far beyond simple correlation measures.

The research underscores the critical importance of:

- Advanced statistical techniques
- Comprehensive market behavior analysis
- Recognizing the multidimensional nature of market interactions

5 Alternative volatility measures

5.1 Introduction

Volatility is a fundamental concept in financial markets, reflecting the risk associated with asset returns. Such measures are critical to portfolio management, derivatives pricing, and risk management. This essay explores the alternative volatility models, focusing on realised volatility and GARCH- based modelling, with reference to Blair, Poon, and Taylor (2001) and Andersen and Benzoni (2008), and is interlaced with emphirical analysis using 5-minute high frequency stock price data of JPMorgan (JPMUS).

5.2 Alternative Volatility Measures

The simplest measure of volatility is the historical volatility. This measure calculates the conditional variance (or standard deviation) of asset returns as a historical variance over a fixed rolling window. While easy to compute and interpret, its major downsides are that, as it assumes constant historical variance obtained within the window for the future volatility, it fails to capture the clustering and persistence of volatility that is observed in real markets. To overcome these setbacks and obtain more accurate prediction of volatility, different types of volatility measures are proposed to this date. The major alternative models are outlined below.

GARCH-based Volatility The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, introduced by Bollerslev (1986), captures volatility clustering and persistence by modelling conditional variance as a function of past squared returns and past variances. The standard form of this class of volatility is given as follows:

$$r_t = \mu + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0,1)$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Although still depending on the same frequency of historical data, this class of models has the advantage that it can flexibly describe features of volatility, such as higher significance of negative returns in volatility, through modification of the functional form.

Implied Volatility This measure is derived from option prices, and it is seen as a forward looking measure which reflects market expectations about future volatility. As a major limitation of this model, it relies on the option pricing model assumptions and may be influenced by risk premiums and market imperfections. Nevertheless, Blair et al (2001) indicated that by carefully adressing these points, implied volatility could provide generally better estimate over the historical volatility and ARCH-type models which also considers intra-day data, for S&P100 index returns in the horizons of up to 20 trading days.

Realised Volatility The realized volatility is a non-parametric and ex-post volatility measure, typically calculated as the sum of squares of high frequency intraday returns. As an important theoretical background, if we can assume a standard Brownian motion given below as a high frequency price process, the 1-day sum of quadratic variation of the diffusion term can be understood as a ex-post measure of a daily return volatility:

$$dp(t) = \mu(t)dt + \sigma(t)dW(t) \tag{10}$$

The major difference from the prior models is that, as seen from the expression above, it can assume stochastic volatility by design. Hence, for using this measure for a lower frequecy volatility, we have to take quadratic variation of the actual realization of high-frequency returns over the corresponding horizon. In this sense, $\sigma(t)$ of (10) does not directly represent the volatility forecast. Hence, as detailed by Anderson and Benzoni (2008), the conditional variance is indirectly derived as:

$$Var[r(t,k)|\mathcal{F}_{t-k}] \approx [RV(t,k;n)|\mathcal{F}_{t-k}] \approx [QV(t,k;n)|\mathcal{F}_{t-k}]$$

Blair et al. (2001) show that realised volatility based on high frequency returns provides enhanced forecasting power for future volatility. In addition, the paper showed that while it can be argued that implied volatility can be more accurate in forecasts, realised volatility contains critical information that are not captured by option prices alone. In this point, Anderson and Benzoni (2008) also discusses realised volatility and its effectiveness, especially when combined with parametric models for volatility forecasting and jump detection. Both papers argue that to improve accuracy and responsiveness, one should use realized volatility alongside traditional volatility models. Therefore, in the following section we assessed the GARCH model estimated from the realized volatility to examine the claims argued in both papers.

5.3 Empirical Analysis using JPMUS High-frequency Data

To evaluate the practical utility of alternative volatility measures, we analysed eight months of 5-minute return data from JPMorgan (JPMUS). The analysis supports and complements the findings of Blair, Poon, and Taylor (2001) and Andersen and Benzoni (2008), particularly regarding the behaviour of volatility, persistence, and the benefits of high-frequency data.

a) Realised Volatility from 5-minute returns Daily realised volatility was computed using 5-minute intraday return data. This high-frequency approach allows accurate tracking of daily volatility spikes and aligns with the literature, which shows realised volatility's efficiency and reactivity to market events.

b) GARCH(1,1) Volatility Estimate from Daily Returns A GARCH(1,1) model was estimated using daily return data. The model captured conditional heteroskedasticity and demonstrated volatility clustering, with moderate persistence consistent with empirical evidence.

The main results are shown in Figure 1. Interpretation of OxMetrics diagnostic plots are as follows:

Interpretation of OxMetrics Diagnostic Plots

- **Density Plot:** Strong kurtosis is observed, highlighting fat tails in return distributions. This supports the use of GARCH models to capture non-normal return behaviour.
- **ACF of Returns:** Weak autocorrelation in raw returns suggests market efficiency, in line with random walk theory.
- **ACF and PACF of Realised Volatility:** Long memory behaviour is visible, indicating persistent volatility—consistent with the arguments of Andersen and Benzoni (2008).
- **Realised Volatility Time Series:** Volatility spikes are evident during periods of market events. Despite large movements, returns remain mean-reverting and centred around zero.

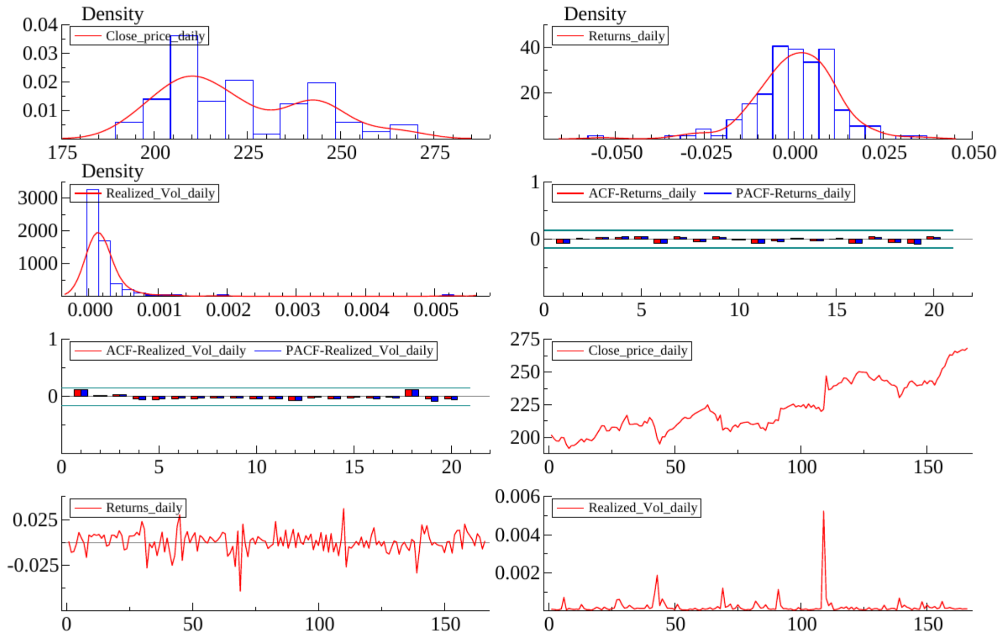

Figure 1: OxMetrics Diagnostic Plots: Density, ACF, PACF and Volatility Structure

GARCH(1,1) Estimation Output from OxMetrics

Parameter	Coefficient	Std. Error	Robust-SE	t-value	p-value
Constant (μ)	0.000894836	0.0007851	0.0007721	1.16	0.248
ω (alpha_0)	4.19861e-05	2.127e-05	2.155e-05	1.95	0.053
α_1	0.243783	0.1187	0.1410	1.73	0.086
β_1	0.441711	0.2110	0.1563	2.83	0.005

Table 10: GARCH(1,1) Model Estimation Results

Summary of interpretation of GARCH(1,1) estimation are as follows:

- **ARCH Coefficient** ($\alpha_1 = 0.2438$): Captures the immediate impact of return shocks on volatility. Suggests that recent squared returns influence current conditional variance.
- **GARCH Coefficient** ($\beta_1 = 0.4417$): Statistically significant ($p = 0.005$), indicating strong persistence of volatility over time.
- **Sum of Coefficients** ($\alpha_1 + \beta_1 = 0.6855$): Since the value is below 1, the model satisfies the stationarity condition. Volatility shocks decay rather than persist indefinitely—suggesting a relatively stable period for the stock.
- **Convergence and Reliability:** The model exhibits strong convergence using the BFGS optimisation method. Robust standard errors support the reliability of parameter estimates.

Conclusion

In summary, this analysis demonstrates the importance of using both realised and model-based volatility measures. As argued by Blair et al. (2001) and Anderson Benzoni (2008), relaised volatility provides incredible sight into actual market variability. On the other hand, GARCH models show insight into volatility persistence. In our empirical analysis of the high-frequency data using JPMorgans returns, the two measures followed similar trends but diverged sharply during volatile periods. In conclusion, both offer a comprehensive framework for forecasting financial volatility.