

---

## Empirical Finance: Assignment 1

---

**By Group 15**

Lanlan Hou	2801069
Rick van der Ploeg	2782774
Sebastiaan van Dijen	2769505
Wa Chak Sou	2796840

School of Business and Economics

**3<sup>rd</sup> November 2025**

## QUESTION 1: Stock prices and returns

### a. Prices and returns plots of HSBC stocks

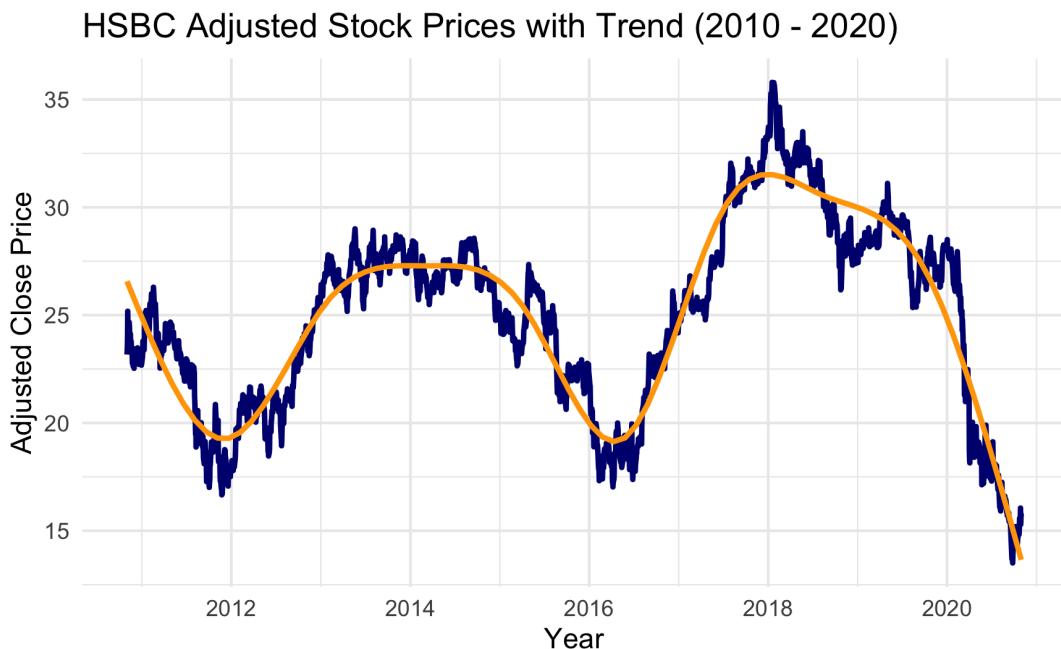


Figure 1.1 Adjusted Close Price of HSBC (November 2010 to November 2020)

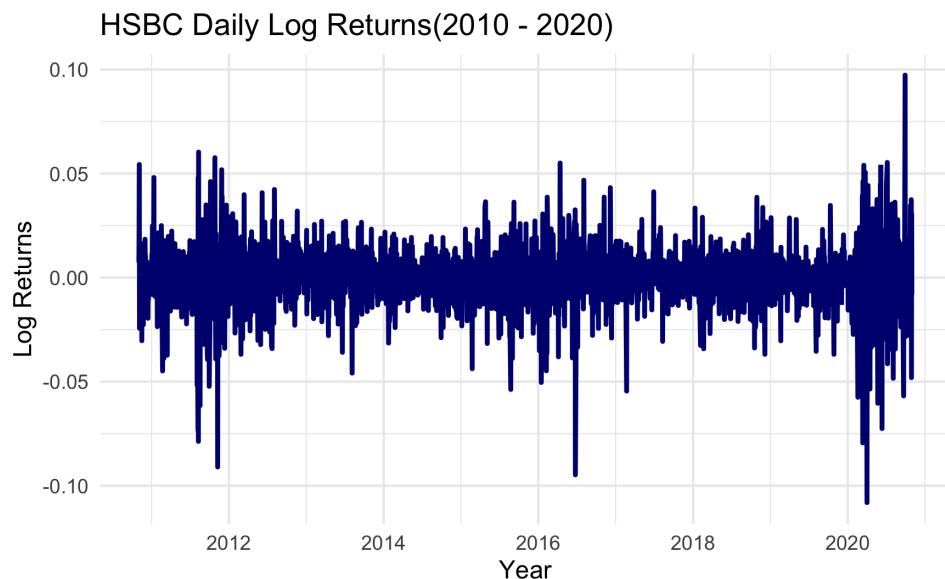


Figure 1.2 Log Returns of HSBC (November 2010 to November 2020)

[Figure 1.1](#) demonstrate the daily stock price of HSBC which fluctuates over the decade. It show a cyclical patterns rather than a clear long-term upward or downward trend. It also reflects the market condition like COVID-19 shock at the end of the decade. [Figure 1.2](#) reveals the daily log returns of HSBC stock which fluctuate around 0.1. The high volatility near 2012, 2016 and 2020 while other periods is low volatility, supports the exhibition of volatility clustering with high and low volatilities interchanging.

### b. Relevant summary statistics

Statistic	Value
Mean	24.994385
Std. Deviation	4.356907
Minimum	13.508404
P10	18.641351
Median	25.693600
P90	30.548970
Maximum	35.793488
Skewness	-0.200543
Kurtosis	-0.553233

Table 1.2.1: Summary Statistics for Daily HSBC Adjusted Close Price

Statistic	Value
Mean	-0.000156
Std. Deviation	0.014615
Minimum	-0.108101
P10	-0.015906
Median	0.000000
P90	0.015354
Maximum	0.097266
Skewness	-0.551902
Kurtosis	8.797429
Autocorrelation (one lag) of returns	-0.049996
Autocorrelation (one lag) of squared returns	-0.051176

Table 1.2.2: Summary Statistics for Daily HSBC Log Returns

From [Table 1.2.1](#) and [1.2.2](#) we observe that the HSBC's adjusted prices fluctuated between roughly \$13.50 and \$35.79 during 2010-2020, with an average around \$25, suggesting moderate long-term volatility(4.36). The daily log returns show a slightly negative mean and moderate volatility. The negative skewness(-0.55) and high kurtosis(8.8) indicate that the return distribution is left-skewness and leptokurtic - implying more large drops than large gains, and extreme movements are more common than in a normal distribution.

## Question 2: Return distributions

### a. Distribution of the returns

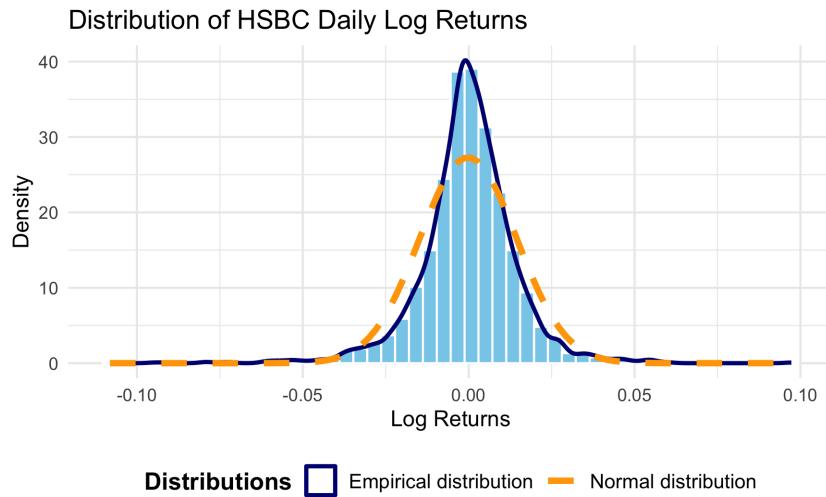


Figure 2.1 Distribution of log returns)

Figure 2.1 shows the distribution of the log returns for HSBC. The histogram of log returns (in *skyblue bar*) is overlaid with the corresponding normal distribution curve (in *orange dashed line*) for comparison. It is evident that the distribution is leptokurtic (more peaked) and exhibits minor skewness, indicating only small deviations from a perfectly normal shape. Overall, the returns appear to be approximately normally distributed, as the normal curve fits the histogram, with no strong evidence of fat tails or extreme outliers.

### b. Normality test of returns

To formally test for normality, the **Jarque-Bera (JB)** test was conducted. The hypotheses for the test are formulated as follows:

$$\begin{aligned} H_0 &: \text{The returns are normally distributed,} \\ H_1 &: \text{The returns are not normally distributed.} \end{aligned}$$

The Jarque-Bera (JB) test statistic, based on the least squares residuals, is given by

$$JB = \frac{T-k}{6}(\hat{S}^2 + \frac{1}{4}(\hat{K}-3)^2),$$

where

- $k$ : the number of covariates (regressors),
- $T$ : the sample size,
- $\hat{S}$ : the sample skewness,
- $\hat{K}$ : the sample kurtosis.

Under the null hypothesis of normality, the test statistic is asymptotically chi-squared with two degrees of freedom:

$$JB \xrightarrow{d} \chi_2^2 \quad \text{as } T \rightarrow \infty.$$

JB Test Statistic	Critical Value (5% level)	p-Value
3202.7	5.99	0.00

Table 2.1: Jarque-Bera (JB) normality test for HSBC Log Returns

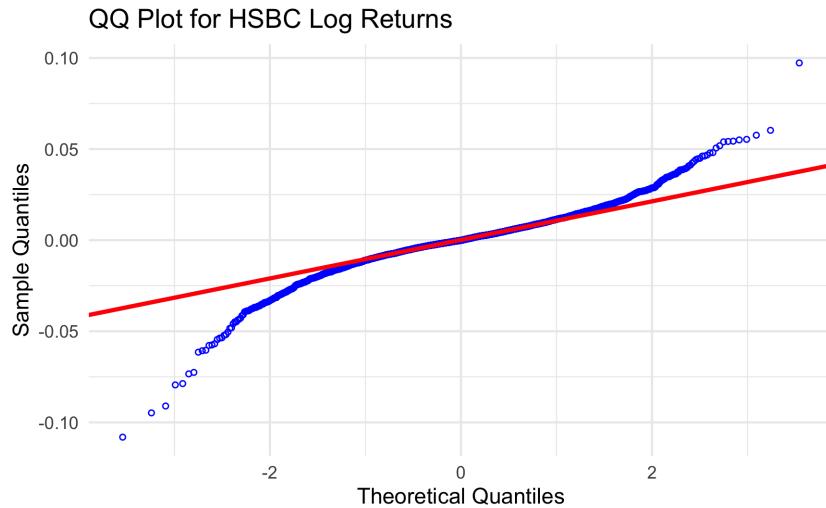


Figure 2.2 QQ Plot for HSBC Log Returns)

**Table 2.1** reports the results of the JB normality test. The computed JB statistic (3202.7) is far greater than the 5% critical value (5.99), and the associated *p*-value is 0.00. Hence, we **reject the null hypothesis** of normality at the 5% significance level. The QQ plot shown in [Figure 2.2](#) further supports this conclusion, showing clear deviations from the theoretical normal line, particularly in the tails.

**Conclusion:** The results of both the Jarque-Bera test and QQ plot clearly indicate that the HSBC log returns are not normally distributed. The distribution exhibits excess kurtosis and fat tails, confirming the presence of non-normal behavior commonly observed in empirical financial data.

## Appendix

```

1 library(openxlsx)
2 raw <- read.xlsx('student_groups_stocks.xlsx', sheet = 1)
3
4 groupNumber <- 15
5 nameOfStock <- raw$Stock.Name[groupNumber]
6 startDate <- raw$Start.Date[groupNumber]
7 endDate <- raw$'End.Date.(+10y)'[groupNumber]
8
9 # Extract the daily stock price from source
10 library(quantmod)
11 getSymbols(nameOfStock, src = 'yahoo', from = startDate, to = endDate)
12 date <- index(HSBC)
13
14 # Adjusted Close
15 adjustedPrice <- as.numeric(HSBC[, 'HSBC.Adjusted'])
16
17 # Log Returns
18 compoundedReturns <- numeric(length(adjustedPrice) - 1)
19 for(i in 2:length(adjustedPrice)){
20   compoundedReturns[i-1] <- log(adjustedPrice[i] / adjustedPrice[i-1])
21 }
22
23 library(ggplot2)
24 prices_df <- data.frame(date = date, adjustedPrice = adjustedPrice)
25 returns_df <- data.frame(date = date[2:length(date)], compoundedReturns
26                           = compoundedReturns)
27
28 # Plot of daily stock price
29 ggplot(prices_df, aes(x = date, y = adjustedPrice)) +
30   geom_line(color = 'navy', linewidth = 1) +
31   geom_smooth(color = 'orange', se = FALSE) +
32   labs(
33     title = paste(nameOfStock, 'Adjusted Stock Prices with Trend (2010
34                  - 2020)'),
35     x = 'Year',
36     y = 'Adjusted Close Price') +
37   theme_minimal()
38 ggsave('figures/prices_plot.png', dpi = 300)
39
40 # Plot of daily log returns
41 ggplot(returns_df, aes(x = date, y = compoundedReturns)) +
42   geom_line(color = 'navy', na.rm = TRUE, linewidth = 1) +
43   labs(
44     title = paste(nameOfStock, 'Daily Log Returns (2010 - 2020)'),
45     x = 'Year',
46     y = 'Log Returns') +
47   theme_minimal()
48 ggsave('figures/returns_plot.png', dpi = 300)

```

Figure 1: Question 1(a)

```

1 # Compute some relevant summary statistics of Adjusted Close
2 library(moments)
3 price_mean <- mean(adjustedPrice)
4 price_stdev <- sd(adjustedPrice)
5 price_minimum <- min(adjustedPrice)
6 price_maximum <- max(adjustedPrice)

```

```

7 price_median <- median(adjustedPrice)
8 price_10_percentile <- quantile(adjustedPrice, 0.1)
9 price_90_percentile <- quantile(adjustedPrice, 0.9)
10 price_skewness <- skewness(adjustedPrice)
11 price_kurtosis <- kurtosis(adjustedPrice)

12
13 price_summaryStatistics <- data.frame(
14   Statistics = c('Mean', 'Std. Deviation', 'Minimum', 'P10', 'Median',
15   'P90', 'Maximum', 'Skewness', 'Kurtosis'),
16   Values = round(c(price_mean, price_stdev, price_minimum, price_10_
17   percentile, price_median, price_90_percentile, price_maximum,
18   price_skewness, price_kurtosis), 6))
19
20 price_summaryStatistics
21
22 # Compute some relevant summary statistics of Log Returns
23 logReturns_mean <- mean(compoundedReturns)
24 logReturns_stdev <- sd(compoundedReturns)
25 logReturns_minimum <- min(compoundedReturns)
26 logReturns_maximum <- max(compoundedReturns)
27 logReturns_median <- median(compoundedReturns)
28 logReturns_10_percentile <- quantile(compoundedReturns, 0.1)
29 logReturns_90_percentile <- quantile(compoundedReturns, 0.9)
30 logReturns_skewness <- skewness(compoundedReturns)
31 logReturns_kurtosis <- kurtosis(compoundedReturns)
32 logReturns_acf_Returns_oneLag <- acf(compoundedReturns, lag.max = 1,
33   plot = FALSE)$acf[2]
34 logReturns_acf_squaredReturns_oneLag <- acf(compoundedReturns^2, lag.
35   max = 1, plot = FALSE)$acf[2]
36
37 logReturns_summaryStatistics <- data.frame(
38   Statistics = c('Mean', 'Std. Deviation', 'Minimum', 'P10', 'Median',
39   'P90', 'Maximum', 'Skewness', 'Kurtosis', 'Autocorrelation (one
40   lag) of returns', 'Autocorrelation (one lag) of squared returns'),
41   Values = round(c(logReturns_mean, logReturns_stdev, logReturns_
42   minimum, logReturns_10_percentile, logReturns_median, logReturns_
43   90_percentile, logReturns_maximum, logReturns_skewness, logReturns_
44   kurtosis, logReturns_acf_Returns_oneLag, logReturns_acf_
45   squaredReturns_oneLag), 6))

```

Figure 2: Question 1(b)

```

1 single_logReturns_df <- data.frame(compoundedReturns =
2   compoundedReturns)
3
4 # Distribution Plot of daily log returns with equivalent normal
5 # distribution curve
6 ggplot(single_logReturns_df, aes(x = compoundedReturns)) +
7   geom_histogram(aes(y = ..density..),
8     bins = 50, fill = 'skyblue', color = 'white') +
9   geom_density(aes(color = 'Empirical distribution'), size = 1) +
10  stat_function(fun = dnorm,
11    args = list(mean = mean(single_logReturns_df$,
12      compoundedReturns, na.rm = TRUE),
13      sd = sd(single_logReturns_df$,
14      compoundedReturns, na.rm = TRUE)),
15    aes(color = 'Normal distribution'),
16    linewidth = 1.5, linetype = 'dashed') +
17    scale_color_manual(

```

```

14   name = 'Distributions',
15   values = c('Empirical distribution' = 'navy',
16             'Normal distribution' = 'orange')) +
17   labs(title = 'Distribution of HSBC Daily Log Returns',
18         x = 'Log Returns', y = 'Density') +
19   theme_minimal() +
20   theme(
21     legend.position = 'bottom',
22     legend.title = element_text(size = 12, face = 'bold'),
23     legend.text = element_text(size = 10))
24 ggsave('figures/returns_distribution.png', dpi = 300)

```

Figure 3: Question 2(a)

```

1 # Perform Jarque-Bera normality test
2 library(tseries)
3 jb_test <- jarque.bera.test(compoundedReturns)
4 jb_test
5
6 # Show the results of the JB test
7 jbTest_summaryStatistics <- data.frame(
8   JB_Test_Statistic = round(jb_test$statistic, 2),
9   Critical_Value = round(qchisq(0.95, 2), 2),
10  p_Value = round(jb_test$p.value, 2)
11 )
12
13 jbTest_summaryStatistics
14
15 # QQ Plot of daily log returns
16 ggplot(single_logReturns_df, aes(sample = compoundedReturns)) +
17   stat_qq(shape = 1, color = 'blue', size = 1) +
18   stat_qq_line(color = 'red', linewidth = 1) +
19   labs(title = 'QQ Plot for HSBC Log Returns',
20         x = 'Theoretical Quantiles', y = 'Sample Quantiles') +
21   theme_minimal()
22 ggsave('figures/returns_qqplot.png', dpi = 300)

```

Figure 4: Question 2(b)