COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Analysis of Five Financial Stocks During the Fed Rate Hike Cycle (2022.02-2023.06)

Project Report

Team 30

Chen Xi, cx2320

Yi Gao, yg2977

Chengcheng Cai, cc5054

IEOR 4150 Introduction to Probability and Statistics

Presented to: Professor A. B. Dieker

December 8th. 2023

1 Background

It's been a blast from the past for Fed rate-setting, with inflation returning as the No. 1 economic threat in the aftermath of the coronavirus crisis. Since the announcement of the first rate hike on March 16, 2022, the Federal Reserve has raised interest rates a total of 10 times by June 2023. During this period, three major U.S. banks—Silicon Valley Bank, Signature Bank, and First Republic Bank—succumbed to consecutive failures, with deposit sizes of $175 billion, $110 billion, and $103.9 billion, respectively.

Given the backdrop of the Fed's interest rate hikes and the concurrent banking crisis, our team aim to delve into the performance of financial stocks and explore potential correlations between them.

2 Description of Data Set

The dataset comprises five stocks from the financial sector for analysis, namely Bank of America (BAC), BlackRock (BLK), J.P. Morgan (JPM), Kohlberg Kravis Roberts (KKR), and Nasdaq (NDAQ). The time range of dataset is from February 1, 2022, to June 30, 2023, which aligns with the U.S. interest rate hike cycle.

3 Illustration of the Functionality of Jupter Notebook

3.1 sigle_analysis. display_hist()

This function is designed to visualize the log return histograms for individual stocks. The bin width is calculated using the formula *Bin Width = (Max - Min) / sqrt(n)*. Additionally, Kernel Density Estimation (KDE) and Normal Distribution Probability Density Function are plotted to enhance the understanding of each stock's log return distribution.

```
# 1. display hist
sigle_analysis.display_hist()
```

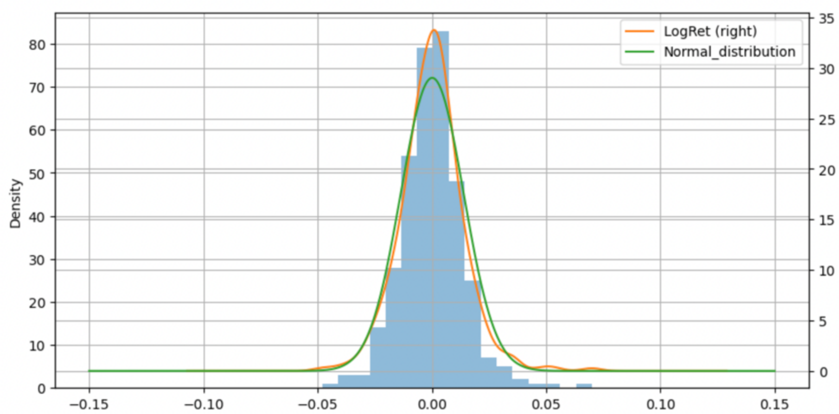

Figure 1 Example of histogram (JPM)

This example displays the histogram of JPM. According to the graph, the log return of JPM during this period roughly conforms to a normal distribution.

3.2 sigle_analysis. display_qqplot()

This function generates Quantile-Quantile (Q-Q) plots for the log-return of each stock using Statsmodels. The purpose is to assess whether the log-return distribution of each stock during the specified period closely approximates a normal distribution.

```
# 2. qq plot
sigle_analysis.display_qqplot()
```

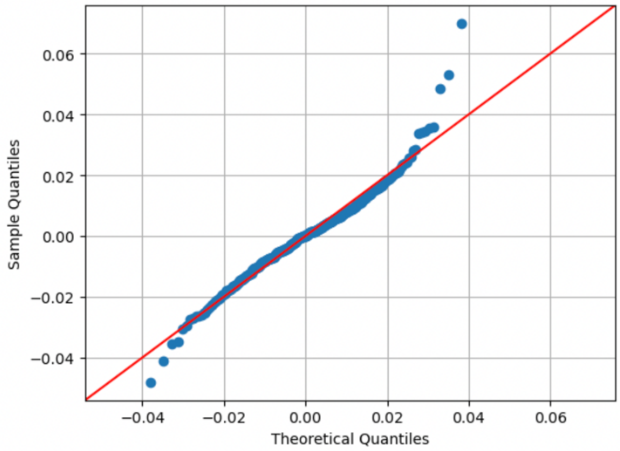

Figure 2 Example of Q-Q plot (JPM)

This example shows the Q-Q plot of JPM. The plot indicates that log returns are roughly normal around the mean but deviate at the tails.

3.3 sigle_analysis. confidence_interval(confidence_level)

This function is used to calculate specific stock's confidence interval of means and variance. Given that the population means and population variance of each stock are unknown, the following methods are utilized to derive the results.

$$100(1-\alpha)\%\ confidence\ interval\ for\ \mu:\ \left[\overline{x} \pm t_{\alpha/2,n-1}\sqrt{\frac{S^2}{n}}\right]$$

$$100(1-\alpha)\%\ confidence\ interval\ for\ \sigma^2:\ \left[\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}\right]$$

```
# 3. confidence interval
sigle_analysis.confidence_interval(confidential_level=0.90)
```

```
90.0% confidence interval for mean:
(-0.0010739112992680618, 0.001332947378682335)
90.0% confidence interval for variance:
(0.0001677670383066989, 0.00021489405945461252)
```

Figure 3 Example of confidence interval of means and variance (JPM)

This example displays the confidence interval of means and variance of JPM based on the 90% confidence level. According to the result, with 90% confidence, the true mean of the JPM is between -0.001073 and 0.001333, and the true variance of the JPM lies between 0.000168 and 0.000215.

3.4 sigle_analysis. regress_on_time()

This function is designed to perform a simple linear regression analysis between time and the log returns of a specific stock. The primary method utilized is Ordinary Least Squares (OLS),

aiming to $minimize\ \sum_{i=1}^{n}\left(y_i - (mx_i + b)\right)^2$.

```
# 4. regression on time
sigle_analysis.regress_on_time()
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                 LogRet   R-squared:                       0.002
Model:                            OLS   Adj. R-squared:                 -0.001
Method:                 Least Squares   F-statistic:                    0.5675
Date:                Tue, 05 Dec 2023   Prob (F-statistic):              0.452
Time:                        02:23:03   Log-Likelihood:                 1018.9
No. Observations:                 355   AIC:                            -2034.
Df Residuals:                     353   BIC:                            -2026.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.0008      0.001     -0.563      0.574      -0.004       0.002
Time_id     5.367e-06   7.12e-06      0.753      0.452   -8.65e-06    1.94e-05
==============================================================================
Omnibus:                       40.168   Durbin-Watson:                   2.052
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              125.449
Skew:                           0.470   Prob(JB):                     5.74e-28
Kurtosis:                       5.757   Cond. No.                         408.
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

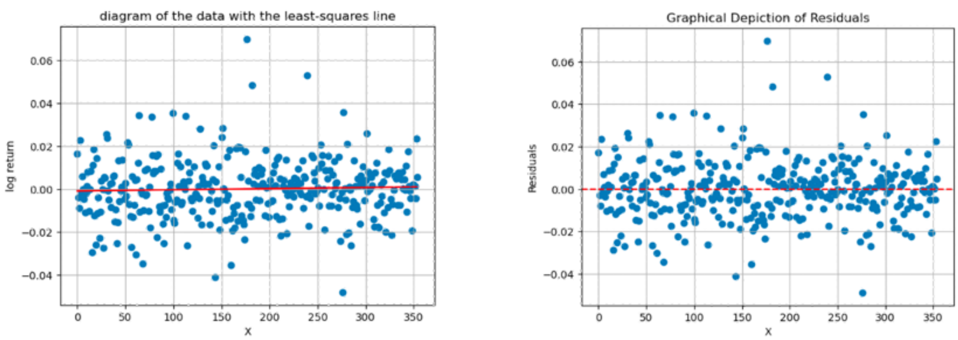

Figure 4 Example of the log-return on time (JPM)

This example displays the simple linear regression analysis between time and the log returns of JPM. According to the table and graph, we can know that there is no significant linear regression correlation between time and log return of JPM.

- R-squared is 0.002, indicating that the model fails to explain the variation in log returns effectively.
- The intercept is -0.0008 with a standard error of 0.001 and is not statistically significant (p-value: 0.574); The slope is 5.367e-06 with a standard error of 7.12e-06 and is also not statistically significant (p-value: 0.452).

3.5 pair_analysis. ttest()

This function is used to test the equality of two specific stocks' population means. Since the prices of two stocks belonging to the same sector may have a degree of interconnection, they are not independent from each other. In addition, n(Stock1) = n(Stock2). Thus, a correlated t-test is employed to examine the equality of the two population means.

Suppose stock1 is X_i, stock2 is Y_i, we set $W_i = X_i - Y_i$

H0: $\mu_w = 0$

H1: $\mu_w \neq 0$

$$T = \frac{\bar{W} - \mu}{S/\sqrt{n}}$$

$$\mu = \bar{x} - \bar{y}$$

$$S = \sqrt{{s_x}^2 + {s_y}^2 - 2rs_xs_y}$$

$if\ T \in \left[-t_{1-\alpha/2,(n-1)}, t_{1-\alpha/2,(n-1)}\right]$, then do not reject H0

if $T \in \left[-\infty, -t_{1-\alpha/2,(n-1)}\right] \cup \left[t_{1-\alpha/2,(n-1)}, +\infty\right]$, then reject H0

```
# t test for equality of mean
pair_analysis.ttest()
```

```
statistic:-1.491057902509486, pvalue:0.13683681006944562
```

Figure 5 Example of correlated t test (BAC and NDAQ)

This example presents the results of testing the equality of population means between BAC and NDAQ. The obtained p-value is 0.1368, which exceeds the significance level of 0.05. Consequently, there is no significant average difference between the log return means of BAC and NDAQ.

3.6 pair_analysis. regression_on_another()

This function is used to conduct a linear regression analysis between the log returns of two specific stocks. It utilized Ordinary Least Squares (OLS) Regression method.

```
In [19]: # regression
         analysis_2.regression_on_another()
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:               LogRet_2   R-squared:                       0.265
Model:                            OLS   Adj. R-squared:                  0.263
Method:                 Least Squares   F-statistic:                     127.4
Date:                Tue, 05 Dec 2023   Prob (F-statistic):           1.94e-25
Time:                        10:55:01   Log-Likelihood:                 1067.6
No. Observations:                 355   AIC:                            -2131.
Df Residuals:                     353   BIC:                            -2123.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0006      0.001      0.972      0.332      -0.001       0.002
LogRet_1       0.4556      0.040     11.289      0.000       0.376       0.535
==============================================================================
Omnibus:                       16.081   Durbin-Watson:                   2.103
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               38.876
Skew:                          -0.091   Prob(JB):                     3.62e-09
Kurtosis:                       4.611   Cond. No.                         63.4
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Figure 6 Example of regression analysis of one stock on the other (BAC and NDAQ)

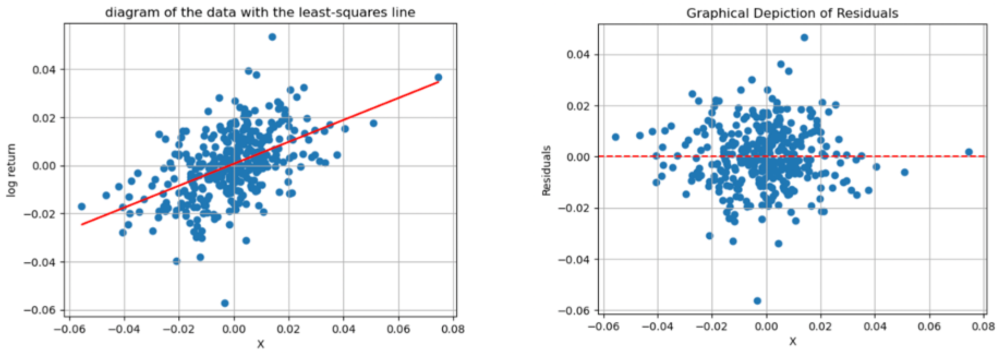

Figure 7 Diagram of regression result (BAC and NDAQ)

This example shows the relationship between BAC and NDAQ. According to the table, the BAC and NDAQ relationship is statistically significant ($p < 0.05$), with an R-squared of 0.265, intercept of 0.0006, slope of 0.4556. And the positive correlation between the two stocks is displayed in the graph.

4 Conclusion

5 Reference

[1] Foster, S. (n.d.). *Fed's interest rate history: The Fed funds rate since 1981*. Bankrate. https://www.bankrate.com/banking/federal-reserve/history-of-federal-funds-rate/