QF603 Group Mini-Project 2

Group F

November 8, 2018

Abstract

In this report, we implemented a simple linear regression of the Dow Jones Industrial Average ("DJIA") index over the S&P500 ("S&P") using daily and annual log returns.

Our key finding is that by using different time intervals (daily vs annual log returns) to perform the regression leads to significant differences in the kurtosis measure of the distribution of the error terms in the regression. Investigating further, the reason for this difference is found to be due to the bias introduced into the annual data series when the last day of the year is used as a benchmark. We present further comparative illustrations to show this.

Task 3: Regression of Daily Log Returns

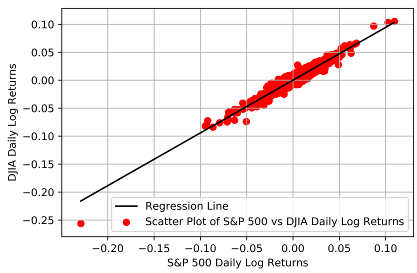

Figure 1: Scatter Plot and Regression Line of Daily Log Returns

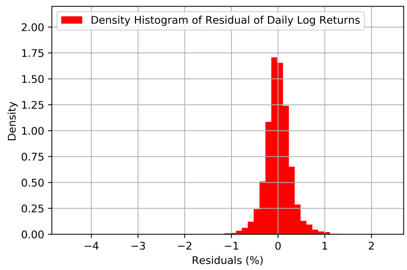

Figure 2: Density Histogram of Residuals

Estimate of key statistics $\hat{a}$, $\hat{b}$ and $\hat{\sigma}_u$

Results

- Alpha $= \hat{a} = 0.00005$, Beta $= \hat{b} = 0.94311$
- Standard Deviation of Residuals $= \hat{\sigma}_u = 0.00288$

The regression can be expressed as $r_{(\mathrm{DJIA},t)} = 0.00005 + 0.94311 r_{(\mathrm{S\&P500},t)}$.

The alpha, or the intercept coefficient, is a small positive number. If alpha is not statistically different from 0, that means that the intercept of the best fit line is at the origin.

The slope of 0.94311 is the DJIA's is the DJIA's Beta - the sensitivity of DJIA's returns to the S&P's returns. In this model the S&P is assumed to be the market, therefore the Beta can be construed as the level of systematic risk.

T-tests for Null Hypotheses a=0 and b=0 at 5% Significance

Results

- T-statistics for $\hat{a} = 1.48636$, and for $\hat{b} = 339.97366$
- Degree of Freedom $= 8500 - 2 = 8498$
- Null Hypotheses (H_0) are that $a = 0 and b = 0$.
- Alternative hypotheses (H_1) are that $a \neq 0 \quad \text{and} \quad b \neq 0$
- Critical value at 5% significance level $= \pm 1.96024$

The test statistic for $\hat{a}$ falls within the critical values, and thus we cannot reject the null hypothesis that $\hat{a} = 0$. This indicates that the DJIA having higher daily log returns than the S&P is therefore not significantly different from 0 at the 5% level.

However, the test statistic for $\hat{b}$ falls outside the critical values, and thus we reject the null hypothesis that $\hat{b} = 0$ and conclude that there is a linear relationship between DJIA and S&P at the 5% significance level.

Goodness of Fit: R^2 and Adjusted R^2 Values

Results

- R-Squared (R^2) = 93.1512%
- Adjusted R-Squared (adj-R^2) = 93.1539%

The R^2 value reports the degree to which our independent variable (the S&P500's daily log returns) explains the variation of the dependent variable (the DJIA's daily log returns). Since R^2 is 93.1512%, it means that over 90% of the variation in the dependent variable is explained by the independent variable.

Adj-R^2 also measures the goodness of model fitting, but takes into consideration the number of independent variables in the model. R^2 will only increase or stay the same when we add more independent variables, even if they do not have any relationship with the dependent variable. Adj-R^2 on the other hand, will "penalise" the model for having excessive dependent variables that do not significantly improve the model.

The adj-R^2 is always lower than the R^2 value. And, in our case, because there is only 1 independent variable, the adj-R^2 and the R^2 values are relatively equal.

Jarque-Bera Test Statistic for the Residuals

Results

- Jarque-Bera test stats = 25434.27
- Degrees of Freedom = 2
- Null Hypotheses H_0: $\hat{\gamma} = 0$, $\hat{\kappa} = 3$
- Alternative Hypotheses H_1: $\hat{\gamma} \neq 0$, $\hat{\kappa} \neq 3$
- Critical Chi-Square Value at 5% significance level = 5.99146

The JB test statistic exceeds the critical value by a huge margin, strongly indicating that the residuals are not normally distributed. This is due to regression outliers that were a result of extreme market conditions, for example the huge one-day drop on 19-Oct-1987.

Additional Test-H_0: $\hat{b} = 1$

<u>Results</u>

- The t-test statistic for $H_0 : \hat{b} = 1$ is -20.50787
- Degree of Freedom$= 8500 - 2 = 8498$
- Null Hypotheses (H_0) is that $\hat{b} = 1$
- Alternative hypotheses (H_1) is that $\hat{b} \neq 1$
- Critical value at 5% significance level $= \pm 1.96024$

We conduct an additional test of $H_0 : \hat{b} = 1$ at the 5% significance level to determine if our $\hat{b}$ value of 0.94311 is statistically different from 1. We do this to gauge how well the returns of the DJIA mimic the returns of the market portfolio S&P.

We reject the null hypothesis that $\hat{b} = 1$, and conclude that the daily log returns of the DJIA do not perfectly mimic the returns of the S&P. Therefore, using the DJIA as a benchmark for performance evaluation may lead to different results as compared to when using the S&P as the benchmark.

Task 4: Regression of Yearly Log Returns

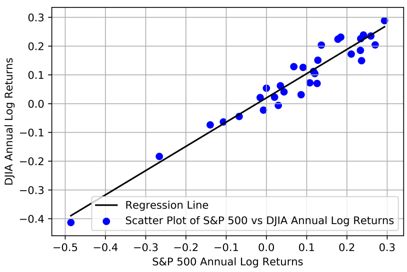

Figure 3: Scatter Plot and Regression Line of Daily Log Returns

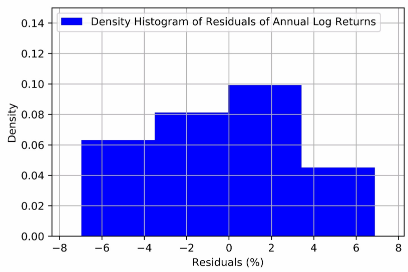

Figure 4: Density Histogram of Residuals

Estimate of Key Statistics $\hat{a}$, $\hat{b}$ and $\hat{\sigma_u}$

<u>Results</u>

- Alpha $= \hat{a} = 0.01978$, Beta $= \hat{b} = 0.84254$
- Standard Deviation of Residuals $= \hat{\sigma}_u = 0.03797$

The regression can be expressed as $r_{(\text{DJIA},t)} = 0.01978 + 0.84254 r_{(\text{S\&P500},t)}$.

The alpha, or the intercept coefficient, is a small positive number. If alpha is not statistically different from 0, that means that the intercept of the best fit line is at the origin.

The slope of 0.8425 is the DJIA's Beta - the sensitivity of DJIA's returns to the S&P's returns. In this model the S&P is assumed to be the market, therefore the Beta can be construed as the level of systematic risk.

T-tests for Null Hypotheses a=0 and b=0 at 5% Significance

Results

- T-statistics for $\hat{a} = 2.64989$, and for $\hat{b} = 20.4364$
- Degrees of Freedom $= 32 - 2 = 30$
- Null Hypotheses (H_0) are that $a = 0 and b = 0$.
- Alternative hypotheses (H_1) are that $a \neq 0 \quad \text{and} \quad b \neq 0$
- Critical value at 5% significance level = ± 2.042272

The test statistic for $\hat{a}$ falls outside the critical values, and thus we reject the null hypothesis that $\hat{a}$=0. The t-test at 5% significance concludes that the DJIA has higher annual returns as compared to the S&P.

The test statistic for $\hat{b}$ also falls outside the critical values, and thus we reject the null hypothesis that $\hat{b}$=0 and conclude that there is a linear relationship between the annual returns of the DJIA and the S&P at the 5% significance level.

Goodness of Fit: R^2 and Adjusted R^2 Values

Results

- R-Squared (R^2) = 93.2983%
- Adjusted R-Squared (adj-R^2) = 93.0749%

The R^2 value is very high, further supporting our previous point that there exists a linear relationship between the annual returns of the DJIA and the S&P.

Jarque-Bera Test Statistic for the Residuals

<u>Results</u>

- Jarque-Bera test stats = 25434.27
- Degrees of Freedom = 2
- Null Hypotheses H_0: $\hat{\gamma} = 0$, $hat\kappa = 3$
- Alternative Hypotheses H_1: $\hat{\gamma} \neq 0$, $\hat{\kappa} \neq 3$
- Critical Chi-Square Value at 5% significance level = 5.99146

The JB test statistic falls within the critical value, indicating that the regression residuals are normally distributed. We note that this finding is in contrast to the JB test performed for the daily log returns, and would like to investigate further as to the possible reason why this may be so.

When we observe the density histograms (at the top of each section) we can see the stark difference in the distribution of the residuals. Although the histogram of residuals from the daily log returns regression has a more bell-like shape, that distribution fails the JB test. On closer inspection, we note that there are large negative outliers in the the histogram of residuals from the regression of the daily log returns (as can be seen from the range of the x-axis). We, therefore, suspect that it is these large outliers that caused the distribution to fail the JB test; by increasing the kurtosis of the distribution beyond what is tolerable under a normal distribution.

When we plot the top 100 “worst day” (i.e. largest loss) log returns by month, we can see that most of the worst days do not even fall in December, let alone the 31st. Clearly, there is a greater number of outliers “included” in the daily return residuals.

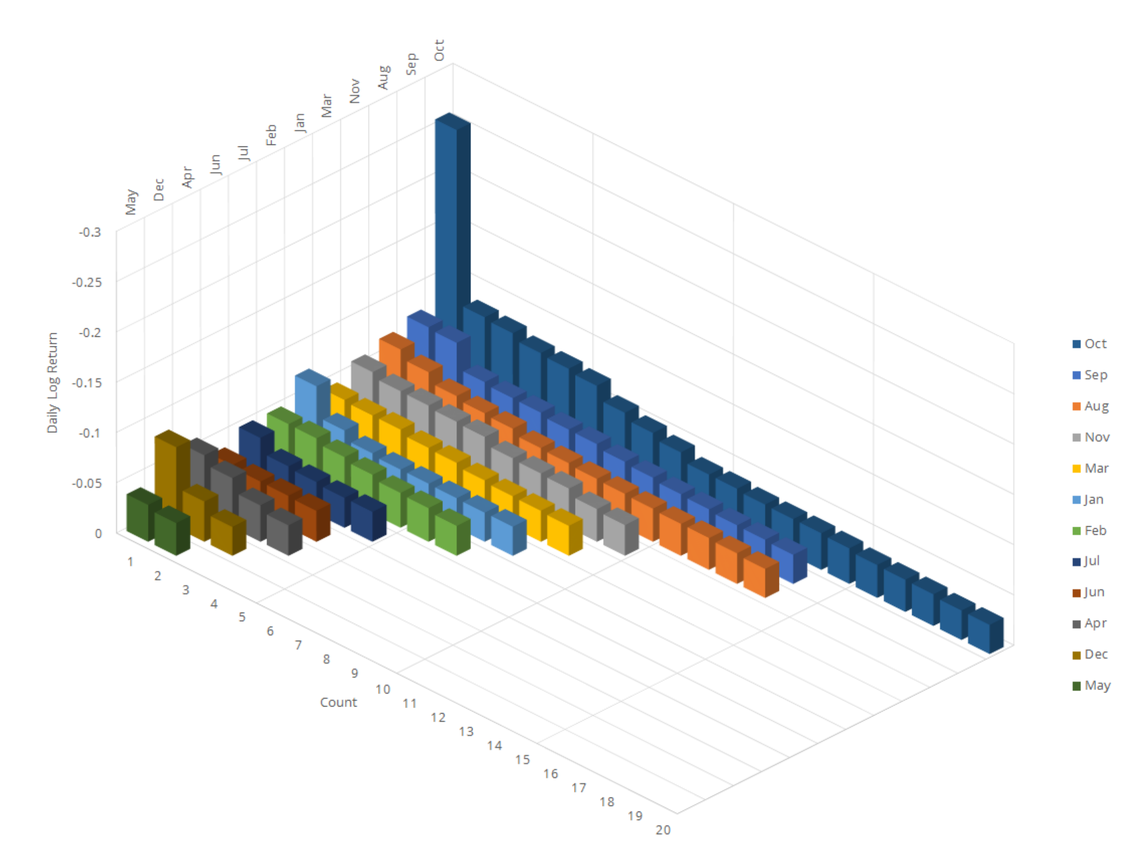

Figure 5: Chart of Top 100 Largest Negative Log Returns

The exclusion of these outliers in the annual log return regression probably "saved" it from failing the JB test.

Additional Test-H_0: $\hat{b} = 1$

<u>Results</u>

- The t-test statistic for $\hat{b} = 1$=-3.81917
- Degree of Freedom=$32 - 2 = 30$
- Null Hypotheses (H_0) is that $\hat{b} = 1$
- Alternative hypotheses (H_1) is that $\hat{b} \neq 1$
- Critical value at 5% significance level = ± 1.96024

We conduct an additional test of $\hat{b} = 1$ at the 5% significance level to determine if our $\hat{b} = 1$ value of 0.84254 is statistically different from 1. We do this to gauge how well the returns of the DJIA mimic the returns of the market portfolio S&P.

We reject the null hypothesis that $\hat{b} = 1$, and conclude that the annual log returns of the DJIA do not perfectly mimic the returns of the S&P. Therefore, using the DJIA as a benchmark for performance evaluation may lead to different results as compared to when using the S&P as the benchmark.