25573 Time Series Econometrics
Coombs 12895809, Zanker 12899389, Whitehill 13215948

25573 Time Series Econometrics Group Assignment

Question 1

Review of the 2002 annual report's for both Vale S.A and Bhp Billiton reveal that a substantial portion of their respective businesses are reliant on the production of iron ore. Due to the similar business operations of the two Company during from 2002 to 2006 they required the same asset base to facilitate operations and further faced the same risks.

As the two companies were heavily dependent on iron ore to generate, revenue and profit, their underlying performance was ultimately determined by the price of iron ore along with some other minor factors which essentially created noise in the two stock returns.

A cointegrating relationship can therefore be observed between the two companies because their revenues were driven by the same underlying factor which ultimately translated to stock returns which moved in tandem over time. The termination of the cointegrating relationship between the two firms is most likely due to the divestment of the two firms interests in iron ore, increasing their dependence on other mineral operations further, the different performance of the two Companies facing the softening market conditions leading into the GFC.

Question 2

The Dickey-Fuller test tests the following hypothesis:

$$H_0: \pi = 0$$
$$H_1 : \pi < 0$$

Where $\pi = (\phi - 1)$ which is from $\Delta z_t = (\phi - 1)_{zt-1} + \epsilon_t =: \pi z_{t-1} + \epsilon_t$.

Process	Test Statistic (tau2)	Critical Value (tau)			Test statistic (phi1)	Critical Value (phi)		
α	-	0.01	0.05	0.10	-	0.01	0.05	0.10
logBHP	-0.3982	-3.43	-2.86	-2.57	1.8698	6.43	4.59	3.78
logVale	-0.2921	-3.43	-2.86	-2.57	2.7093	6.43	4.59	3.78

A cointegrated relationship is defined by two log I(1) processes, and a linear combination process of I(0). For a I(1) process to exist, the result of the Dickey-Fuller null hypothesis must fail to reject.

In the case of the log prices of BHP and Vale, the processes both contain unit roots, or I(1). This is due to the tau2 test-statistic taking on a value greater than the critical values, and the phi1 test-statistic taking on a lesser value than the critical values.

Question 3

The regression model;

$$s_{2,t} = \alpha + \theta s_{1,t} + z_t$$

Is estimated as;

$$s_{2,t} = 1.785049 + 0.753770 s_{1,t} + z_t$$

Coefficients	T Value	T Critical Value	P-Value
Intercept	358.5	1.646121	<2e-16
df$logVale ($s_{1,t}$)	266.6	1.646121	<2e-16

For the estimated parameters to be statistically significant, they have to have a p-value of less than 0.05 (the chosen alpha), and the t-value must exceed the critical value.

In this case, the estimated parameters for the regression model are both highly statistically significant.

Question 4

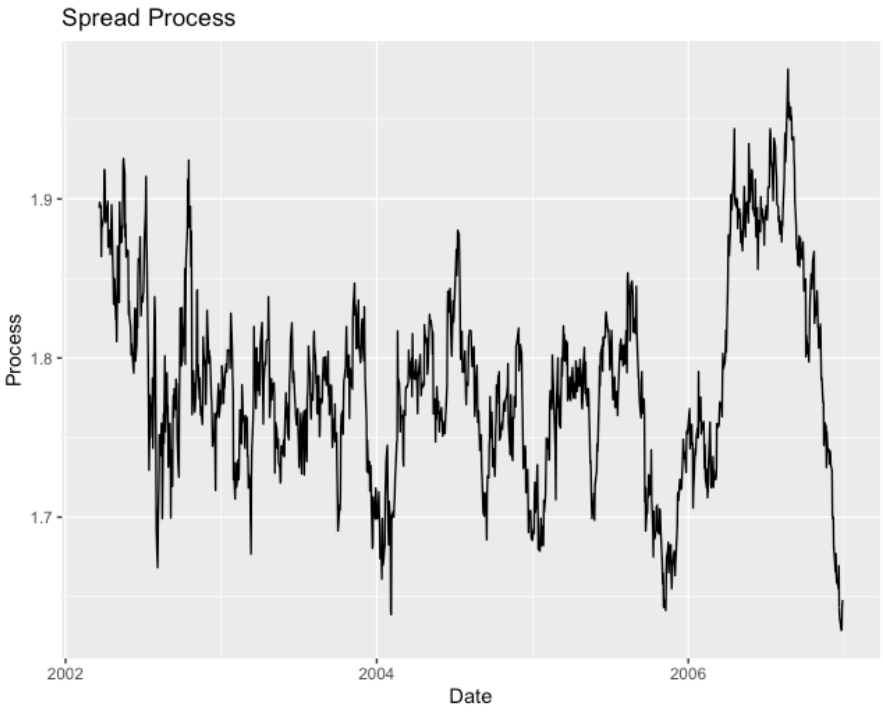
The Dickey-Fuller test tests the following hypothesis:

H_0: $\pi = 0$
H_1 : $\pi < 0$

Process	Test-statistic	Critical Values			
Regression Residuals (z_t)	-4.3404	α	0.01	0.05	0.10
			-2.58	-1.95	-1.62

The test statistic is smaller than all of the critical values, therefore the null hypothesis that the process contains a unit root can be rejected. This means that the residuals process is stationary, or is an I(0) process. This is consistent with a cointegrating relationship between two log-prices because for a cointegration relationship to exist, the log processes must be I(1) (as discovered above) and the residuals process must be I(1) (as discovered here).

Question 5



Question 6

The ECM estimate for the processes is:

$$\Delta s_{2,t} = 0.009594 + 0.1210678 - 0.1506076\Delta s_{2,t-1} - 0.0302007\hat{z}_{t-1} + \upsilon_t$$

The estimated error correction coefficient $\hat{\gamma}$ value is -0.0302007. The p-value of this coefficient is statistically significant, its value is 0.000726. The value of $\hat{\gamma}$ is negative, this allows the cointegration relationship to begin to 'correct' itself with each time step, approaching to the long-term mean

Question 7

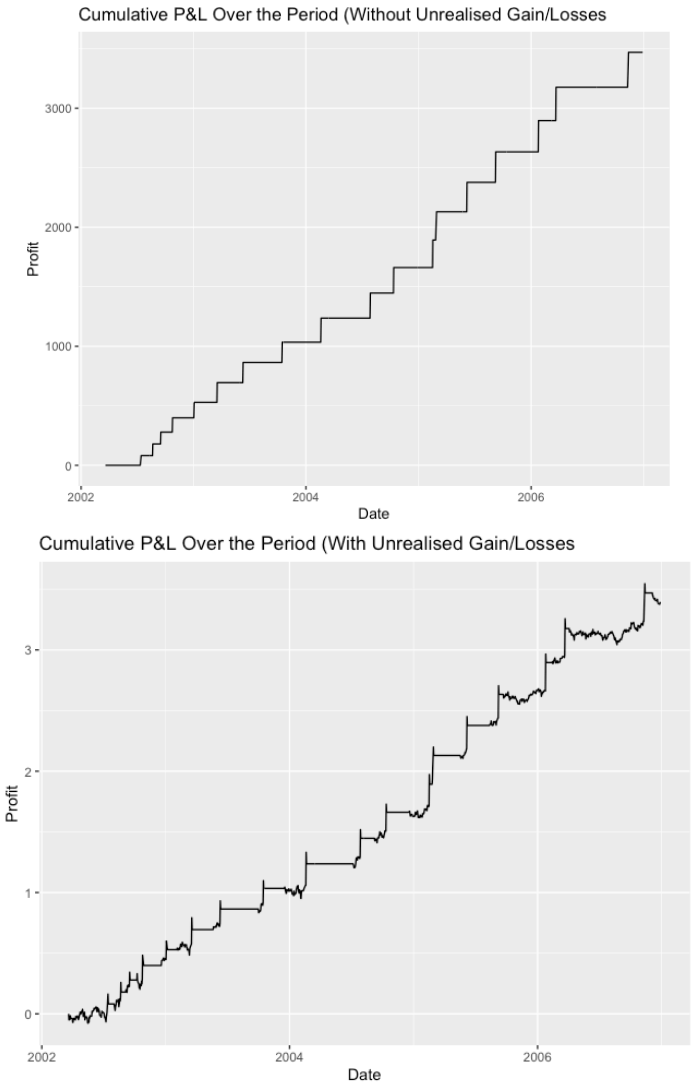
The Johansen test for cointegration yields the results of an error correction coefficient -0.03084. This is close to the value from the two step Engle-Granger method.

According to the first column in the code, which is statistically significant creates the following equation:

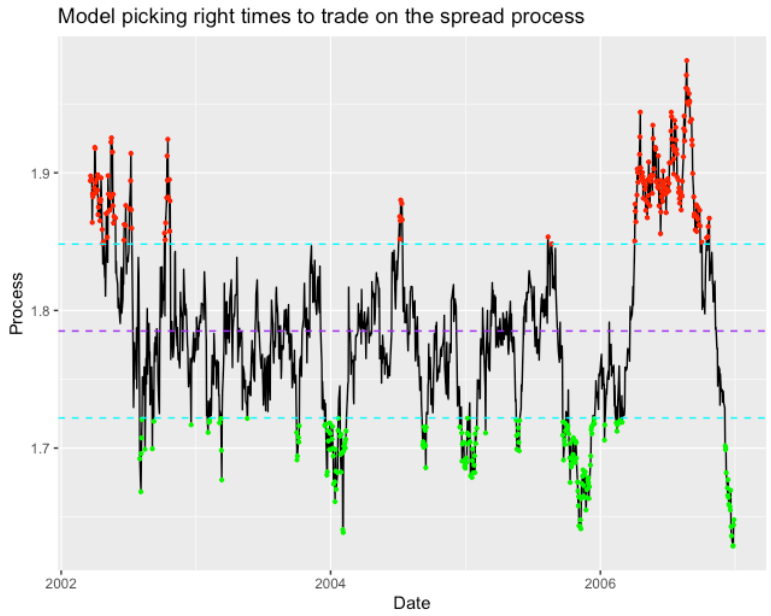
$$y_t - 0.7594961x_t = 1.7793155 + \begin{pmatrix} 0.7594961 & 1 \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix}$$

Theta from the two step Engle-Granger method, was 0.7537705, where the ECM estimates theta at 0.759461. The results from the two methods of testing for cointegration are in agreeance with each other. Therefore, it can be concluded that the two log-stock processes are cointegrated.

Question 8



In addition to the cumulative profit and loss shown above, the ‘buy’ and ‘sell’ positions of the spread process and boundaries. Green is the long and red is the short position.

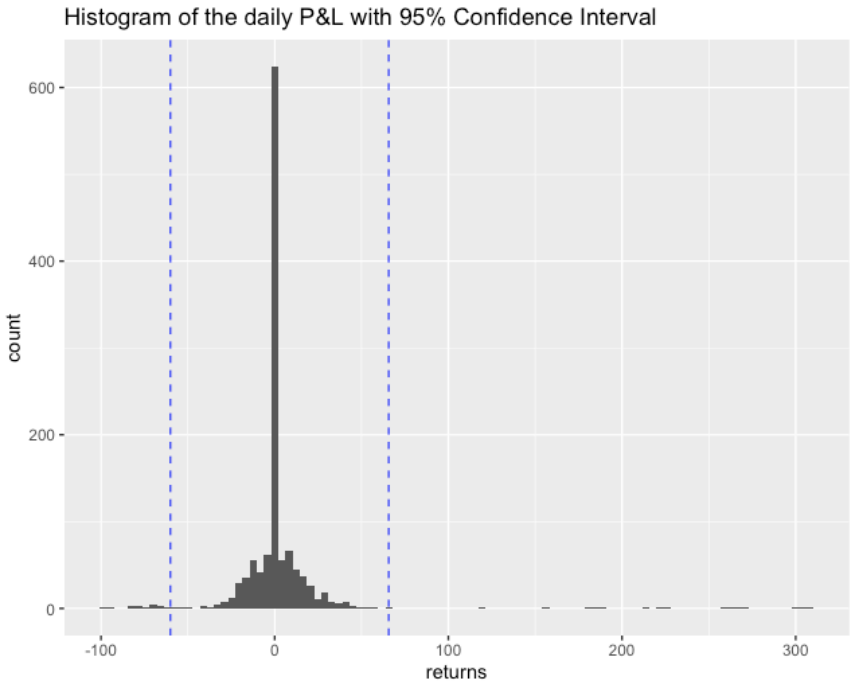


The red is when the spread process is "sold" i.e. short BHP and long theta of Vale.
The green is when the spread process is "bought" i.e. long BHP and short theta of Vale.

Notes on the programming.
A few improvements to the code, and therefore improving the reliability of our long and short positions would be to;

- Have a theta (linear combination ratio) that moves along with time,
- Have a moving w_*, so that the threshold is better represented by the most recent data,
- Have a sampling period because it is evident that the process begins to break down. This is evident around the 2nd quarter of 2006, because the process begins to deviate from its stationary characteristics, which are important to the statistical arbitrage opportunity.

Question 9



The worst daily loss with 95% certainty is -$62.

Question 10

The risk-free rate in the Sharpe ratio has been assumed to be 5%, an average from 2002-2007. Therefore, the Sharpe ratio of the trading strategy is 1.509602.