**Fin 608 Project 2**

Group 2

### Background

Pairs Trading Strategy is a simple arbitrage strategy. The logic behind is that we find two stocks have moved together historically and focus on the spread between them to determine our long and short position when a price diverge is detected, then we simply wait for them to converge again. In this project, we aim to find 5 pairs of stocks among the given data set which contain 66 stocks in oil and gas extraction industry. We would consider all the possible pairs, and calculate their tracking variance, which would be one of our criterions to determine the final pairs. Also, we would focus on fundamental analysis to further check whether the picked pairs are reasonable.

### Identify pairs with lowest tracking variance

The basic pairs screening process can be described as follows.

First, with a quick glimpse of the data set, we found that four companies had some missing data among their returns during the given time window, which are WHD, CEPU, BWP and WGL. Hence, they are excluded from the data set, which left us with 62 stocks available.

Second, we computed the tracking variance of all possible pairs among these 62 stocks, which leads to 1861 outcomes. We ranked all the pairs and picked top 5 pairs with lowest tracking variance, which are CNP & SRE, ATO & OGS, NJR & OGS, NI & OGS and NI & SR.

Finally, we conduct the fundamental analysis on these 5 pairs to further check whether our choices are reasonable or are just by coincidence. If something undesirable was detected on some pair, that pair would be replaced with another pair with a relative higher tracking variance, but more reasonable fundamental information. The fundamental analyses of above pairs are as follows:

**CNP&SRE:** CNP engages in power generation and distribution. SRE operates the developing and operating energy infrastructure and provides gas and electricity services to their customers in North and South America.

Since their business highly matches with each other, CNP and SRE should be a reasonable pair.

**ATO$OGS:** ATO regulates Natural Gas Distribution and pipeline and storage business. OGS operates the provision of Natural Gas Distribution services. Both of their businesses cover the Natural Gas Distribution service. Hence, ATO and OGS should be a reasonable pair.

**NJR&OGS:** NJR provides the distribution of natural gas through a regulated utility, which provides other retail and wholesale energy services to customers and invests in clean energy projects and midstream assets. Both of their businesses contain the segment of Natural Gas Distribution. Hence, NJR and OGS should be a reasonable pair.

**NI&OGS:** NI provides natural gas, electricity, and other products and services. It operates through the following segments: Gas Distribution Operations, Electric Operations, and Corporate and Other. Both of their businesses include the segment of Natural Gas Distribution. Hence, NI and OGS should be a reasonable pair.

**NI&SR:** SR provides natural gas service through its utility operations while engaging in non-regulated activities. Their businesses are highly correlated with each other. Hence, NI and SR should be a reasonable pair.

Besides, for the seven companies above, we extract their fundamental data from FactSet. Here, we use EPS, PE ratio, ROE and Debt Asset Ratio, which indicates a company’s capability of generating profits, price level and solvency.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | EPS | P/E | ROE | D/A |
| CNP | 4.13 | 6.87 | 43.99 | 37.76 |
| SRE | 1.01 | 105.86 | 2.00 | 38.47 |
| ATO | 5.43 | 17.29 | 13.91 | 30.69 |
| OGS | 3.08 | 23.79 | 8.47 | 28.69 |
| NI | 0.39 | 65.82 | 3.07 | 42.39 |
| SR | 3.43 | 21.76 | 8.58 | 39.29 |
| NJR | 1.52 | 27.73 | 10.99 | 36.36 |

Based on the accounting data, we can see that all of them have positive EPS, all PE ratios except for that of company SRE are not too high, ROE and Debt Asset Ratio are in the appropriate range. Thus, these companies are reasonable candidates.

### Trading performance test

During the whole 6-month test period (02/04/2018-26/09/2018), since transaction cost are ignored, basically the arbitrage strategy can make an arbitrary amount of money, therefore in order to uniform the result, we assume a capital on long position of $100,000 and a capital on short position of $100,000 during each trading period, which results in a total capital of $0 for each trade. In other words, every time the deviation of stocks within a pair exceeds 2 sigma, we open our position and assign equal capital of $100,000 on long and short side; and we close our position when the price converge or the test period is over.

Trading performance of single pairs are shown in figure 1 to figure 5, and the total profit and loss of the portfolio consisting these five pairs are shown in figure 6. For each pair, the stock prices are normalized based on their price at March 31, 2017, which is the starting date of the training data. Besides, the position are shown in black lines under each figure, where a higher one represents the stocks are being held and a lower one represents no stocks are being held.

#### 3.1 CNP and SRE

As shown in figure one, the position opened twice during the test period. The first trade locked a profit of $7910, however the second position was forced to close at the end of the test period, though a profit was achieved, it may result in a potential loss in the future if the stocks could not converge as expected.

#### 3.2 ATO and OGS

As shown in figure two, this pair of stocks showed a greater than 2 sigma gap at the beginning, therefore the position was opened at the beginning and closed successfully within the test period, resulted in a profit of $9715. After that, no extra trades were conducted.

#### 3.3 NJR and OGS

As shown in figure three, this pair of stocks showed a quite narrow spread at the beginning. However, two signals for opening with gaps greater than 2 sigma were detected during the test period, though the second one just happened in a flash. At the end of test period, we achieved a total profit of $5588.

#### 3.4 NI and OGS

As shown in figure four, the position was opened only once during the test period. Since their price did not converge after the opening, we were forced to close them at the end of the test period. One can see from the figure that NI experienced a major plunge near the end of the test period, which results in even larger gap at the closed time; therefore we suffered a loss of $8418.

#### 3.5 NI and SR

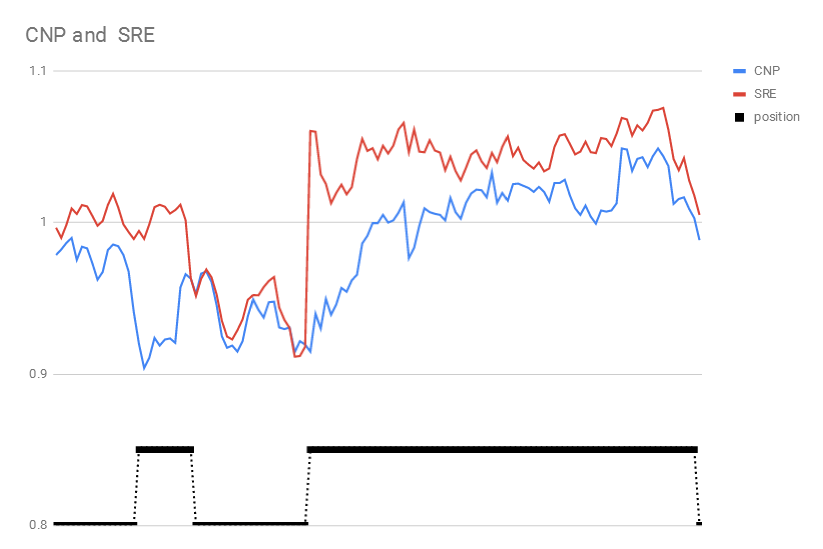
As shown in figure five, the position was opened twice during the test period. The first trade locked a profit of $7073 and the second trade was opened late during the test period so we forced it to close at the end of the test period. Due to the plunge of NI, a relative large amount of profit was achieved in the end.

#### 3.6 Portfolio Performance

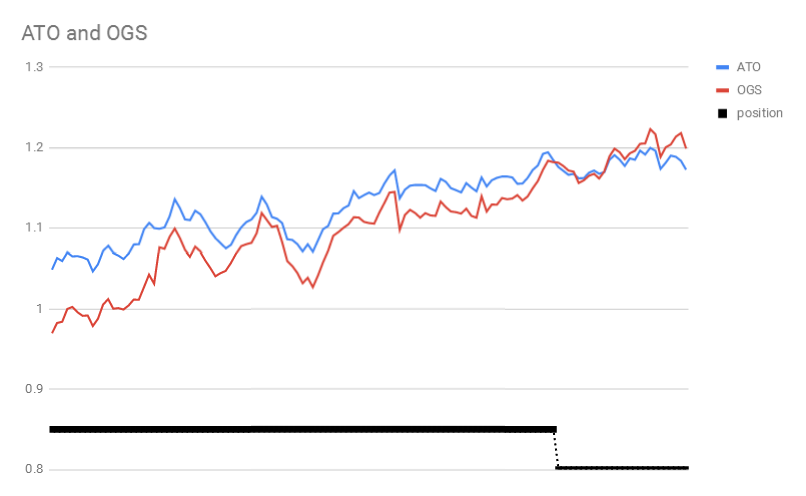
As shown in figure six, by trading on these pairs, the total portfolio value grew steadily during the test period.

### Appendix

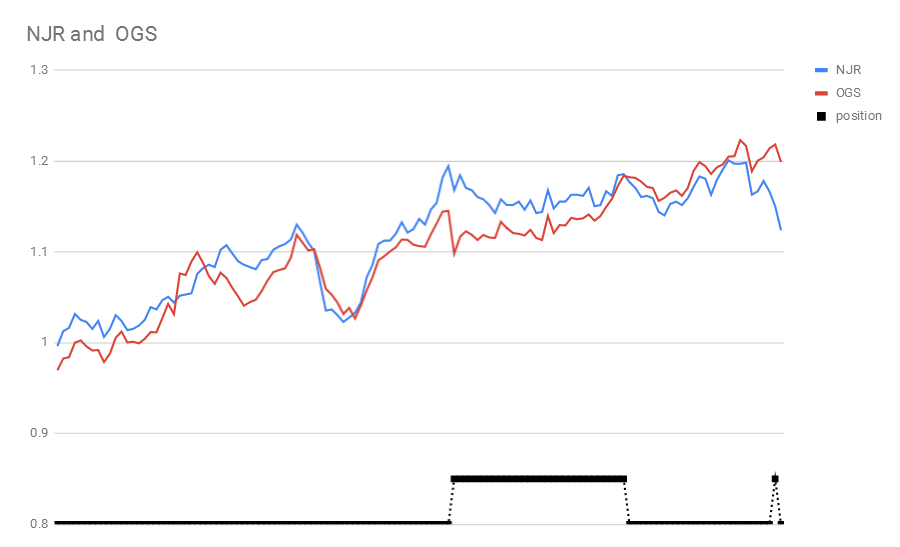
**Figure 1**



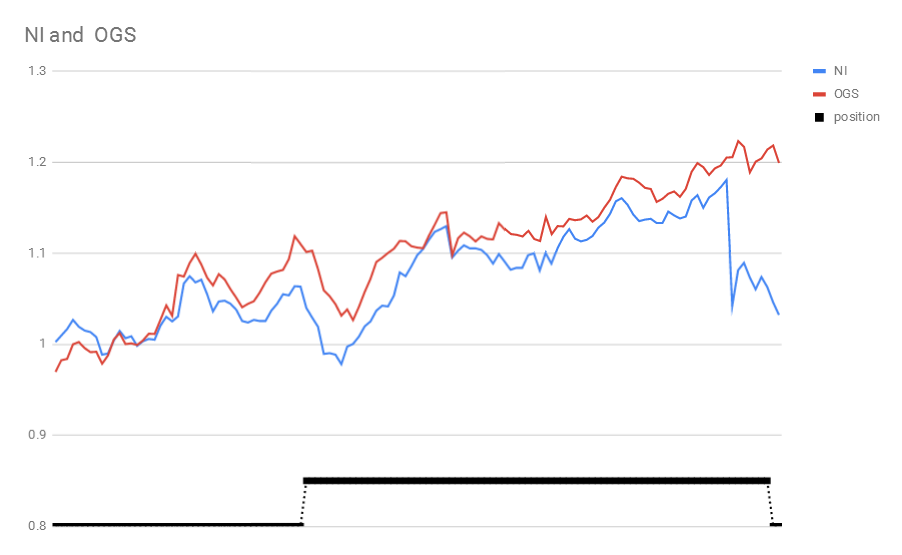
**Figure 2**



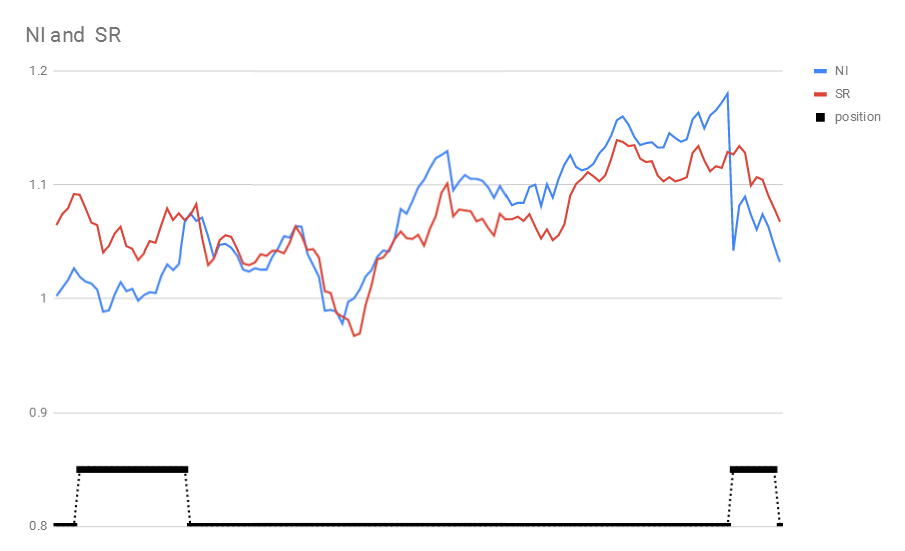
**Figure 3**



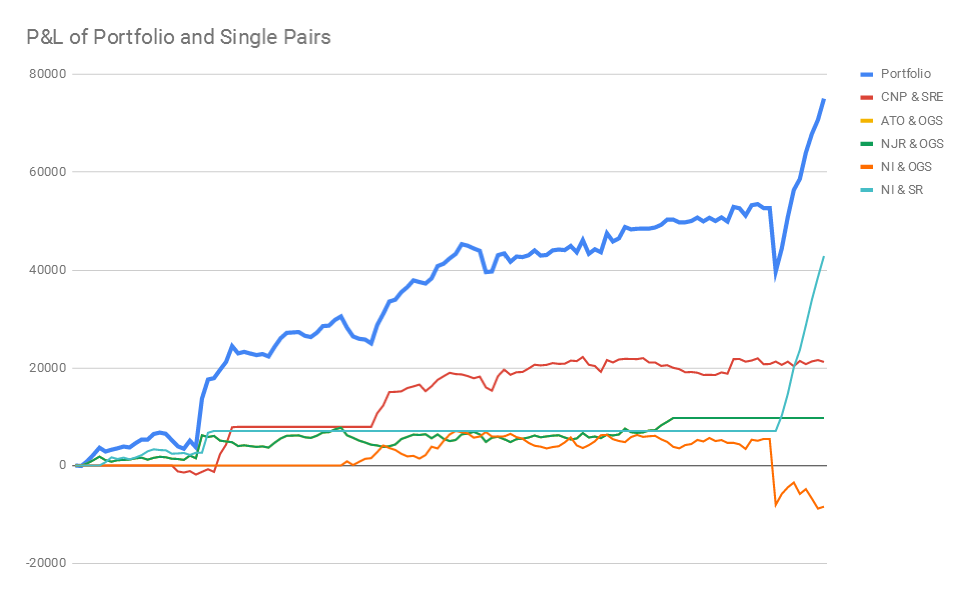
**Figure 4**



**Figure 5**



**Figure 6**



**Code**

We used the following Python code for tracking variance calculation.

import pandas as pd

import numpy as np

import itertools as it

def tracking\_variance(x,y):

price\_x = np.ones((len(x),1))

price\_y = np.ones((len(y),1))

for i in range(1,len(x)):

price\_x[i] = price\_x[i-1] \* (1 + x[i]/100)

price\_y[i] = price\_y[i-1] \* (1 + y[i]/100)

return (((price\_x - price\_y).T).dot(price\_x - price\_y))[0][0]

my\_data = pd.read\_excel("data.xlsx")

my\_data = my\_data.dropna(axis=1)

my\_data.index = [pd.to\_datetime(d) for d in my\_data['TICKER']]

my\_data = my\_data.drop('TICKER',axis=1)

pair\_list = list(it.combinations(my\_data.keys(),2))

tracking\_var\_list = [tracking\_variance(my\_data[\_[0]],my\_data[\_[1]]) for \_ in pair\_list]

min\_five\_pairs = np.argsort(tracking\_var\_list)[0:15]

*#15 paris are reported in case some of them would be excluded*

min\_five\_pairs\_tikers = [pair\_list[i] for i in min\_five\_pairs]

**Google Sheet Visit Link**

Please use the following link to visit our Google sheet for more detailed excel calculation process.

<https://docs.google.com/spreadsheets/d/137Z8LyoLFhozi11M6tJyTuHm7jogkzrxNmpYN9QC2Qk/edit?usp=sharing>