Pairs Trading Simulation Study

Introduction

The economy revolves around finance and financial trading. To maximize profits in buying and selling stock, it is vital to have algorithms and models to predict opportunities and strategies to implement at such moments. Pairs trading is one such method that attempts to automate such a process. Its implementation on tens of thousands simulated time series exhibits that it is most effective when the dependence (cross-correlation of short time lags) between two stocks is high.

Background

Pairs trading is a early model of stock trading that was at the peak of use in the 80s. It established that the ratio of stock prices of similar companies tends to stay consistent. Variations from the mean are then opportunities for profit since the model predicts that the direction of the ratio will tend back towards the mean. Past historical data for the two companies is used to establish the magnitude of variation at which to buy and sell.

Methods

In order to implement the pairs trading strategy, R functions were created to implement the following rules:

- When the ratio of stock 1â€™s price over stock 2â€™s price moves above k standard deviations from the long term mean (i.e., $r > m + ks$), sell \$1 worth of stock 1 and buy \$1 worth of stock 2 (for simplicity, we are assuming we can buy fractions of a share). Then we wait until the ratio is less than or equal to m, at which point, we close the position by buying back however many shares of stock 1 we initially sold, and selling the shares of stock 2 we initially bought.
- Similarly, when the ratio is less than *m âˆ’ ks*, do the same thing but reversing the roles of stock 1 and stock 2. In this case we wait until the ratio rises back up to be greater than or equal to m.

Additionally, in order to optimize k for a given pair of stocks, we trained the strategy on historical values by iterating over a range of values for k from 0.5 to the maximum deviation observed. The theoretical profits were calculated for trading at each value of k and based on \$1 worth of stock bought and sold at each opening position. The k-value which maximized profit for the given historical data was chosen to determine future trading positions.

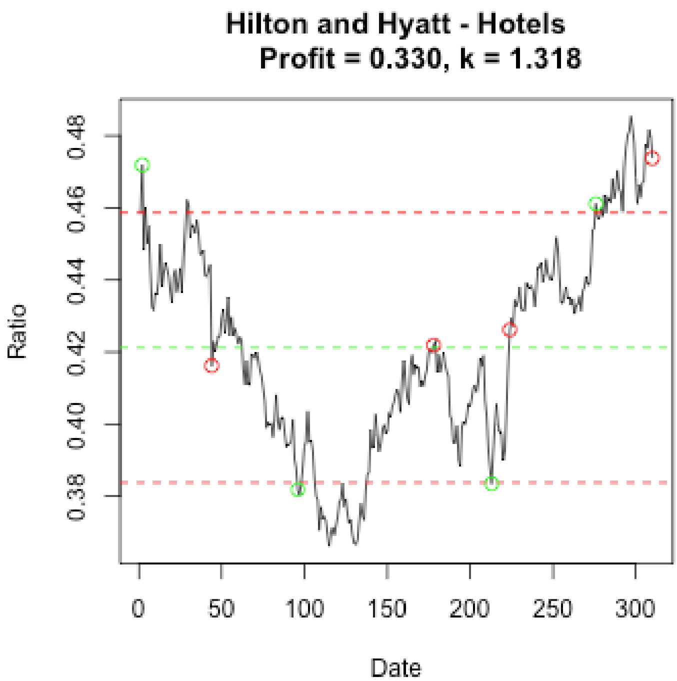
Empirical Study

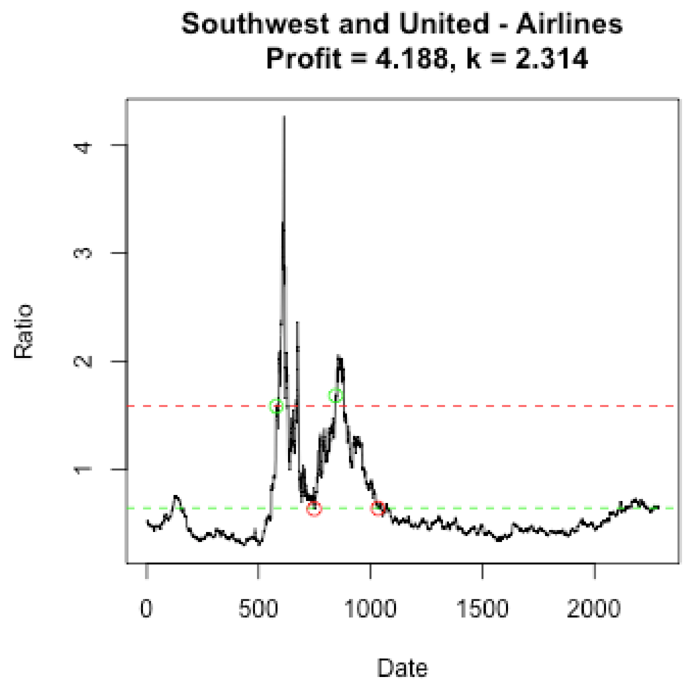
To test the functionality of the pairs trading model created, I compared six pairs of stocks from companies of similar industries. The two companies that are compared need to be of the same industry in order to reasonably make the assumption that the ratio of their stock prices exhibits properties of a weakly stationary process (i.e. the ratio of their stock prices will converge to a central value).

Empirical data is based on daily closing prices of 14 publicly traded companies across various industries (Ford, GM, IBM, Intel, Alcoa, Bayer, Boeing, Hilton, Hyatt, Kellogs, Mariott, ProcterGamble, Southwest, and United).

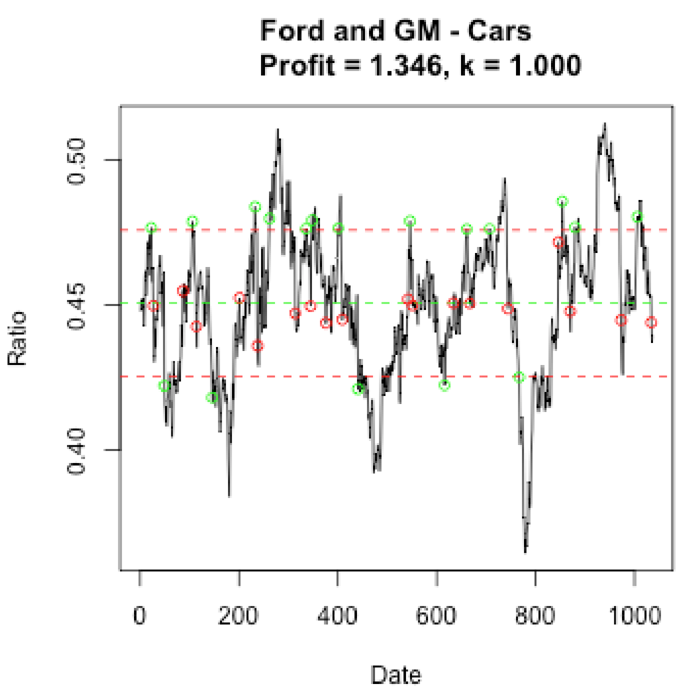
One of the notable trends was that stocks that have mean ratios far from 1 (Hotels - Hilton and Hyatt) tend to have lower profit margins compared to stocks of similar value, i.e. ratio ~ 1 (Airlines - Southwest and United). Additionally, stocks with few trading positions, but peaks of high magnitude, showed greater profits than those with more positions but less deviation (Industry vs Cars).

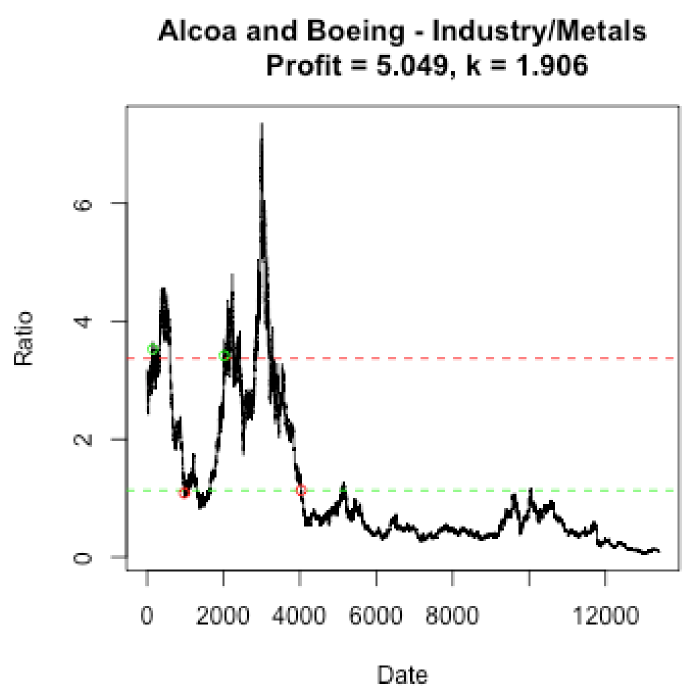
--- Difference in ratio mean





--- Difference in deviation from mean: frequency and magnitude





Simulation Study

Similulations of hypothetical time series were created based on 8 parameters: rho, psi, 2 beta0's, 2 beta1's, and sigma. Each of 300 parameter combinations were used to find a distribution of the profits of 100 time series each of which had 4000 values. To determine approximate parameter values to iterate over, I ran and plotted a few simulations varying each term to get a sense of the effect that each has on the ratio of the stock values, as well as which values are reasonable for simulation. Through this process, the observed effects of each parameter are: rho is inversely proportional to the amount of times the ratio fluctuates and thus increases the frequncy of buy/sell positions, psi is inversely proportional to the sd of the ratio and thus keeps our stocks from deviating to far from the mean, beta1 is proportional to the variance of the corresponding stock, and sigma is the amount of variation introduced when creating the stock values.

After iterating through 300 variations* of parameters for the time series simulation, there is an apparent pattern to the mean values of the resulting profits. First, I sorted the 300 distributions by the rho value that was used to create them from .95 to .99 (60 each) and then took the mean value for set of distribution means. As rho decreases in value, the mean profit tends to increase (see values below) (rho = .99 shows a higher value than expected, probably due to variation in simulation).

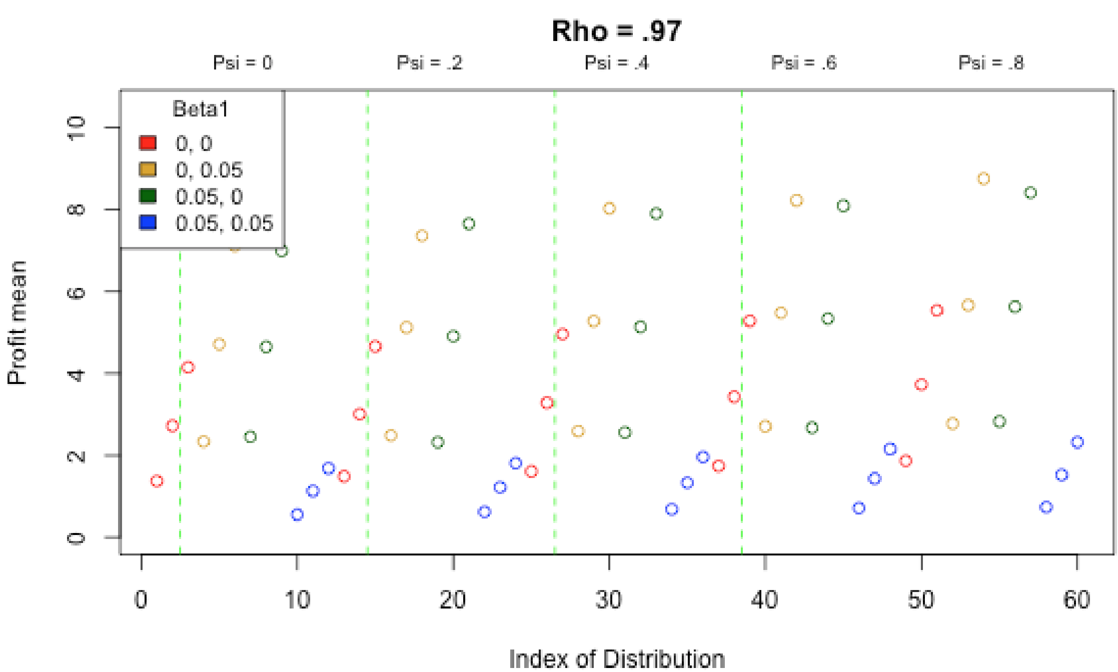
After identifying the trend in rho and looking inside a particular value of rho, there again are apparent trends for each of the parameter values. Separating by value of psi shows a notable pattern (see example below for rho = .97) which shifts slightly higher in profit for psi = 0 to .8.

Finally, going into each psi value, one can see that higher profits are observed when the values for beta1 are not equal, and higher variation sigma for the randomness introduced in our time series. The difference in beta1 is indicative of a pair of stocks for which one has more day to day fluctuation while the other is more consistent which produces a higher profit in our pairs trading model.

Seeing these trends, the optimal parameters for the pairs trading model to maximize profit are lower rho, higher psi, differing beta1's, and higher sigma within the range of reasonable values for each.

*Rho (.95, .96, .97, .97, .99) *Psi (0, .2, .4, .6, .8) *beta1.1 (0, 0.05) *beta1.2 (0, 0.05) *sigma (.5, 1, 1.5)

```
##    Rho ProfitMeans
## 1 0.95    4.558831
## 2 0.96    4.176164
## 3 0.97    3.749538
## 4 0.98    3.317764
## 5 0.99    3.593829
```



Discussion

A primary factor that this model of pairs trading does not take into consideration when calculating profit is the cost of buying and selling. This model predicts higher profit for lower rho values (more buy/sell positions); however, this might be suboptimal as the cost of buying/selling multiple times adds up.

Additionally, our time series simulations assumed a great amount of dependence between the two stocks that keeps the ratio of the two stock prices around a central value, and that such dependence will be consistent from past to present. Although this may be reasonable for pairs of companies in the same industry for most time periods, instances in which deviation from this pattern exist, e.g. one company coming out with a great innovation, could result in a great amount of profit loss. Relevant information about the potential for such instances needs to be introduced in order to implement pairs trading effectively.

Lastly, time constraints and computing power limited the number of test cases and simulations that could be performed. The test was performed assuming that one observes the stock values only once a day, an assumption that misses many profit opportunities if used in the actual stock market.

Conclusions

The model of pairs trading excels when the two stocks under examination exhibit a great dependence on one another. The dependence of two companies drives the ratio of their prices towards a long-run average which in turn effects the number of buy and sell opportunities, which correspond to opportunities to profit, and possibly more importantly, reduces probability of a long-term loss by segmenting risk into short lower risk profit opportunities.