A Possible Way to Search Pairs Trading Arbitrage

Ruming Liu, rumingli@usc.edu

Instructor: Ricardo Mancera, mancera@usc.edu

October 2020

Math 590 Directed Research

Algorithms Trading Practice

Mathematical Finance

Department of Mathematics

University of Southern California

A. Introduction

In traditional statistical arbitrage, pairs trading is a very popular method in industry, which is based on identifying pairs of stocks that typically trade in a predictable relation to one another. Usually their linear combination may have a long run equilibrium like below:

$$log(S_1) = \beta_0 + \beta_1 t + \beta_2 log(S_2) + \epsilon \tag{1}$$

Sometimes, if stock pairs are chosen perfectly, there is no time trend in (1), the model will be simpler:

$$log(S_1) = \beta_0 + \beta_1 log(S_2) + \epsilon \tag{2}$$

We can run the regression of any two stocks 1 & 2. Once the coefficients are significant and error term ϵ is a purely white noise with variance σ_ϵ^2, then if we structure a portfolio V (Long 1 log(stock 1) and short β_1 log(stock 2)). It will be a stationary time series with:

$$E[V(t)] = \beta_0 \tag{3}$$

$$Var[V(t)] = \sigma_\epsilon^2 \tag{4}$$

How can we arbitrage from this property? If the economy is stable, our portfolio V price should be stable with β_0 too. A much higher (or lower) price means you can short (or long) this portfolio to get profit with low risk.

When we select the specific stock 1 & 2, a common way is to find two highly correlated stock in the same industry, such as Walmart with Targets, McDonald with Burger King, etc. We can run the regression of these stock pairs and see whether they have cointegration relationship. If they have, we can generate profits from short run unbalance of their price gap. The motivation of this method is hoping gap is stationary, make profits from volatility. Because the higher volatility means higher trade frequencies.

B. Method for Searching Possible Stock Pairs

Two natural questions are:

Can we have other criteria for choosing stock pairs, not only depends on their correlation?

How to make arbitrage if volatility is low?

We can consider this problem from portfolio dynamic point of view. Suppose for any two stocks 1 & 2, their dynamics are:

$$dS_1(t) = \mu_1 S_1(t)dt + \sigma_1 S_1(t)dW(t) \tag{5}$$

$$dS_2(t) = \mu_2 S_2(t)dt + \sigma_2 S_2(t)dW(t) \tag{6}$$

By Ito formula, we can generate dynamic of portfolio V:

$$dV(t) = \left[\mu_1 - \beta\mu_2 + \frac{\beta\sigma_2^2 - \sigma_1^2}{2}\right]dt + (\sigma_1 - \beta\sigma_2)dW_t \tag{7}$$

If in equation (7), it is a Brownian motion with drift. If the drift term of portfolio V is 0. Then:

$$V(t) = V(0) + \int_0^t (\sigma_1 - \beta\sigma_2)dW_t \tag{8}$$

$$E[V(t)] = V(0) \tag{9}$$

By Ito isometry

$$Var[V(t)] = E\left[\int_0^t (\sigma_1 - \beta\sigma_2)dW_t\right]^2 = E\left[\int_0^t (\sigma_1 - \beta\sigma_2)^2 dt\right] \tag{10}$$

Which is very similar with the model in equation (2), (3), (4).

If in equation (7), there exists positive drift term. Then:

$$V(t) = V(0) + \int_0^t [\mu_1 - \beta\mu_2 + \frac{\beta\sigma_2^2 - \sigma_1^2}{2}]dt + \int_0^t (\sigma_1 - \beta\sigma_2)dW_t \tag{11}$$

$$E[V(t)] = V(0) + [\mu_1 - \beta\mu_2 + \frac{\beta\sigma_2^2 - \sigma_1^2}{2}]t \tag{12}$$

$$Var[V(t)] = E\left[\int_0^t (\sigma_1 - \beta\sigma_2)^2 dt\right] \tag{13}$$

Which is very similar with the model in (1), one interesting thing is that if variance in (13) is small, from (12) we know the price of portfolio V will increase (or decrease) in a stable trend, we can long (or short) portfolio V to make arbitrage.

Instead of finding high correlated company in the same industry and run regression to analyze its cointegration, maybe we can use numerical method to estimate the parameters of stocks dynamic. According parameters to judge the arbitrage.
What's more, in our model stock 1 & stock 2 are still high correlated because they are motivated by a same Brownian motion. If we set different Brownian motions for different stocks, maybe we can find arbitrage opportunities in low correlated stocks.

Now we define a Z score to describe the spread relationship between any two stocks.

$$Z = \frac{\mu_1 - \mu_2 + \frac{\sigma_2^2 - \sigma_1^2}{2}}{\sigma_1 - \sigma_2} \tag{14}$$

Which is very similar with Sharpe ratio of portfolio which is constructed by longing 1 log(stock 1) and shorting 1 log(stock 2).

How does Z score can help us to make decision of choosing stocks pairs? The answer is choosing those stocks pairs whose Z score is low. It is absolutely sure, low Z score means in equation (11), the drift term is small, and volatility term is high. Which is what we want, the spread of log(stock 1) and log(stock 2) is stable in the long run, but in the short run there exists high volatility. Like we said before, the volatile in the short run means sometimes the spread will be too large (small), we can short (long) the portfolio and wait their spread converges to long run equilibrium. Of course, in reality, the portfolio is structed by longing 1 stock 1 and shorting β stock 2, it behaves muck like its log value, so we can use this Z score to choose stock pairs.

C. Pseudocode of Searching Pairs Algorithms

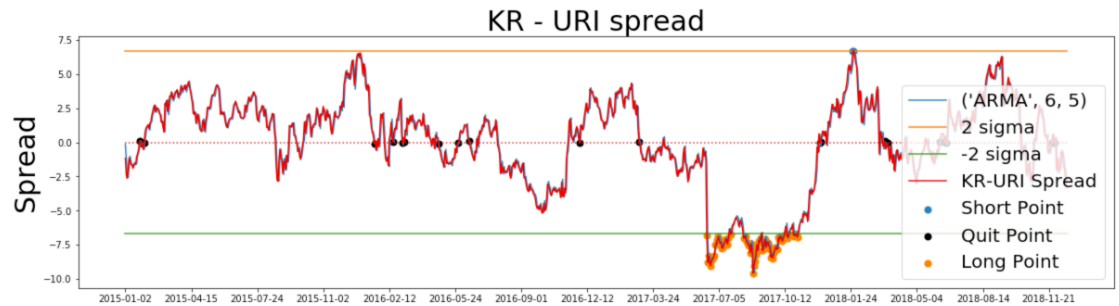
The next part we will use algorithms to find some possible pairs in S&P 500 components according to Z score discrimination. We choose the lowest 50 Z scores pairs to run Time-Series regression and analyze.

First, we will show our pseudocode below:

1. Make all combinations of any 2 stocks in S&P 500 components, and compute Z score of each combination.
2. For each stock i, find a best match stock j, which means their Z score $Z_{i,j}$ is lower than any $Z_{i,k}$,where $k \neq j$.
3. After that, we run the regression $S_1 = \beta_0 + \beta_1 S_2 + \epsilon$ among all pairs we find in step 2.
4. If ϵ in step 3 is stationary, we can say that portfolio structed by longing 1 stock 1 and shorting β stock 2 is stationary. Then we can use ARMA model to estimate and forecast the portfolio value.
5. If ϵ in step 3 is non-stationary, we can check whether it has time trend or it is difference stationary, if yes, we can use detrending technique or ARIMA to estimate and forecast the portfolio value.
6. If one pair doesn't work in step 4 nor step5, we should discard this pair.

D. Time-Series and Cointegration Method

This part, we will show some data examples of those stock pairs which have high Z score and pass the test of step 4 or step 5 and use ARMA model to do analysis and trading.



We use the spread of KR and URI as an example of ARMA method. In this pairs trading, their correlation is -0.712. And by cointegration analysis, our spread portfolio can be constructed by longing 1 KR and longing 0.089 URI. From the graph we can see that spread mean is 0. The spread is weak stationary and we used ARMA(6, 5) model to regress it. We also draw the 2 sigma lines, which indicate that the spread is significant different from its mean 0. When the spread is above the 2-sigma line, it is overvalued, we should short this spread, in the graph we marked as “Short Point”. When it’s under the negative 2-sigma line, it is undervalued, we should long this spread, we marked as “Long Point”. Then we hold the position till the spread converges to 0 and cover our position, we marked as “Quit Position”.

From the graph we can see that, we have 1 long point (appeared in Sep 2017) and 1 short point (appeared in Feb 2018), both situations spent around 4 moths converging to 0.

This method usually needs investor to be patient, they should execute the trading and wait the spread converges to its mean. We may meet horizon risk, which means it spends much long time to converge. For this situation, we can use Optimal Control method to trade because it usually set a terminal time T, what we want is maximizing the wealth at terminal time T.

C. Optimal Control Method

We can also think about this question from stochastic control approach like Jurek and Yang (2007). They assume the spread follows an OU-process and the latter is compounded continuously with the risk-free rate. They set constant relative risk aversion and maximizes the discounted utility of terminal wealth. Utilizing the asset price dynamics, they develop budget constraints and wealth dynamics for the arbitrageurs' assets. Applying stochastic control theory, the authors are able to derive the Hamilton–Jacobi–Bellman (HJB) equation and subsequently find closed-form solutions. We will follow this step to make our trading decision.

The arbitrageur's value function at time t – denoted by V_t takes the form:

$$V_t = supE_t\left[e^{-\beta(T-t)} log(W_T)\right] \tag{15}$$

We allow the arbitrageur to invest in a riskless asset and the mean-reverting spread, their dynamics are below:

$$dB_t = rB_t dt \tag{16}$$

$$dS_t = \kappa(\bar{S} - S_t)dt + \sigma dZ \tag{17}$$

If we denote the number of units of spread and risk-free assts held by N_t and M_t, the portfolio wealth dynamic will be:

$$dW_t = N_t dS_t + M_t dB_t \tag{18}$$

Plug (16) and (17) into (18):

$$dW_t = (r(W_t - N_t S_t) + \kappa(\bar{S} - S_t)N_t)dt + \sigma N_t dZ \tag{19}$$

We can rewrite (15) as:

$$V(W(\tau), S(\tau), \tau) \ = \max_{N(S(\tau),\tau)} \ E_t[e^{-\beta(T-t)} log(W_T)|N(\cdot)] \quad (20)$$

Solving the HJB equation, we can get the optimal control:

$$N = \frac{(\kappa(\bar{S} - S\ ) - rS)}{\sigma^2} W \qquad (21)$$

How can we interpret this result? It's similar with Sharpe ratio for the spread, which is excess return divided by its variance.

Now we come back to our OU process in (17), how can we estimate the parameters of it? We can use the discretized method, we divide time into n discrete time, which gives with $t = k\,\Delta t$:

$$S_{k+1} - S_k = \kappa\bar{S}\Delta t - \kappa S_k\Delta t + \sigma(W_{k+1} - W_k) \qquad (22)$$

Rearrange and set $\sigma(W_{k+1} - W_k) = \sigma\sqrt{\Delta t \epsilon_k}$, we get:

$$S_{k+1} = \kappa\bar{S}\Delta t - \big(\kappa\Delta t - 1\big)S_k + \sigma\sqrt{\Delta t \epsilon_k} \qquad (23)$$

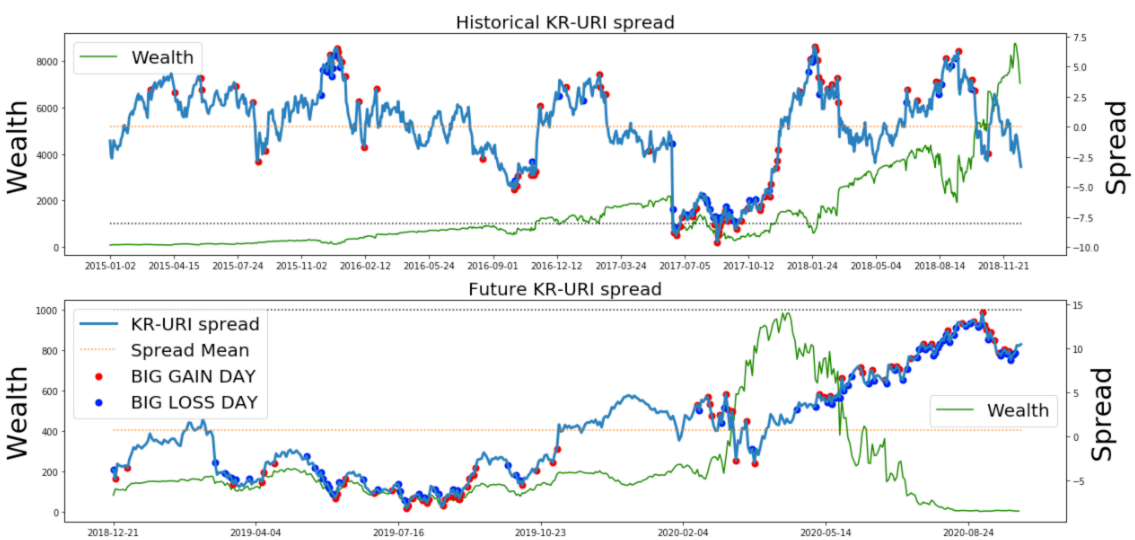
We can use Maximum Likelihood Estimation for this AR(1) process in (23), then we can solve the parameters $(\kappa, \bar{S}, \sigma)$.

Below we will show some data of our pairs trading:

At the beginning, we have \$100 for each pairs trading, we will put our money into spread asset and risk-free asset which we assume risk-free rate is zero. We use the data from 2015-2018 to estimate the parameters and check whether we can make money by optimal control method. If our strategy works, then we will use the data from 2019-2020 to test whether our strategy works out of historical sample.

We will discuss 2 possible scenarios of our strategy:

Scenario 1: Historical data works but future spread diverges

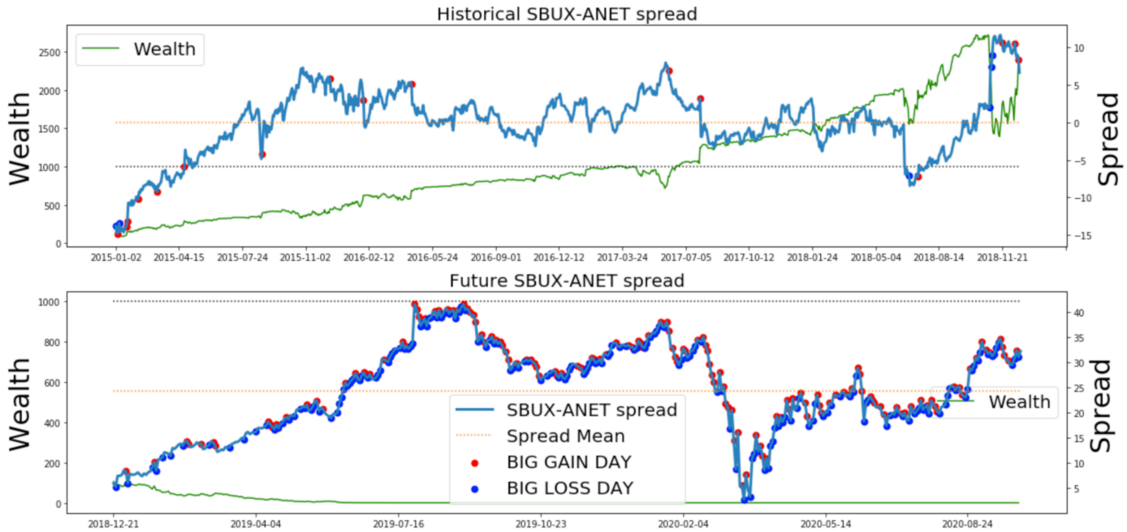


The historical spread of KR & URI is the blue line of first picture. And the red point is the date whose daily return is more than 10%, blue point is the date whose daily return is less than -10%. It is equivalent that red point is high-speed convergent point, blue point is high-speed divergent point. Orange dotted line is the mean of historical spread which is 0. We notice that these two kinds of points don't appear near the orange line, because from equation (21) we know that, when the spread is close to its mean 0, we will not put a lot money into the spread, so our profit is mainly from risk-free asset, daily return will close to 0. Green line is the dynamic of our wealth, we can see it grew from 100 to 8000 during 2015-2018.

The second graph "Future" spread is the data from 209-2020. First, we should notice that it's mean spread is still close to 0, so we can still use the OU process parameters of historical data. Our wealth grew from 100 to 1000 during 2019 to 2020 March. But unfortunately, the Covid-19 pandemic caused the spread divergence, we lose our wealth.

This is the scenario 1, spread diverges during the Covid-19 pandemic.

Scenario 2: Historical data works but future spread mean changes



The first historical graph is the same story with scenario 1. But in second graph, we notice that, the future spread mean changes to 25, but we still use the parameters which assume spread mean is 0. For example, when the spread is 10, our model indicated it will go to 0. However, it actually goes to 25. That's the reason, we lose our initial wealth 100.

These two scenarios told us, even though the historical data works well, the future data may not work so perfectly because its divergence or mean changing. But we can set stop profit and stop loss order. When our wealth for each spread trading is over $1000 or under $10, we will quit this spread trading. In the first graph of scenario 1, we plot a black dotted line of $1000, once our wealth hit the line, we quit, we will earn $1000, which is better than lose all money due to the spread divergence.

If we set our stop order, we trade 7 different spreads, initial wealth is $700. Our final wealth of historical is $7000. Our final wealth of future is $3000, it works well, the stop order prevents us from the Covid-19 pandemic.

Reference:

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=882536

http://folk.ntnu.no/skoge/prost/proceedings/acc08/data/papers/0479.pdf

https://onlinelibrary.wiley.com/doi/abs/10.1111/joes.12153

https://www.sciencedirect.com/science/article/abs/pii/S0167637716302930

https://quant.stackexchange.com/questions/17156/how-can-i-estimate-the-ornstein-uhlenbeck-paramters-of-some-mean-reverting-data

Books:

[Tomas Bjork] Arbitrage Theory in Continuous Time

[James Hamilton] Time Series Analysis

[Andrew Pole] Statistical Arbitrage

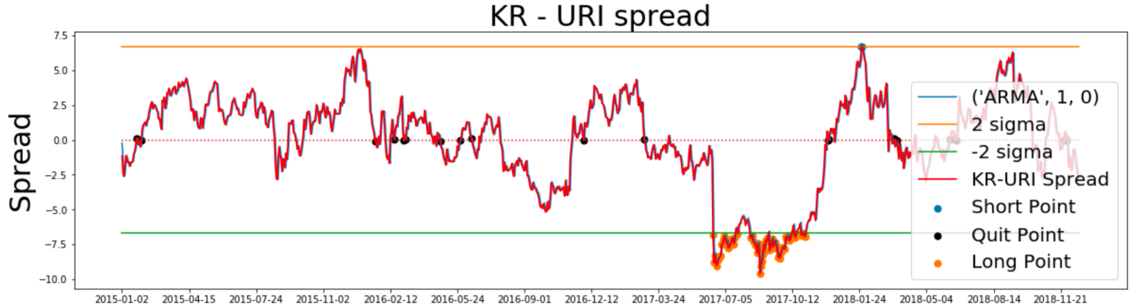
[Walter Enders] Applied Econometric Time Series

[Bodie, Kane, Marcus] Investment

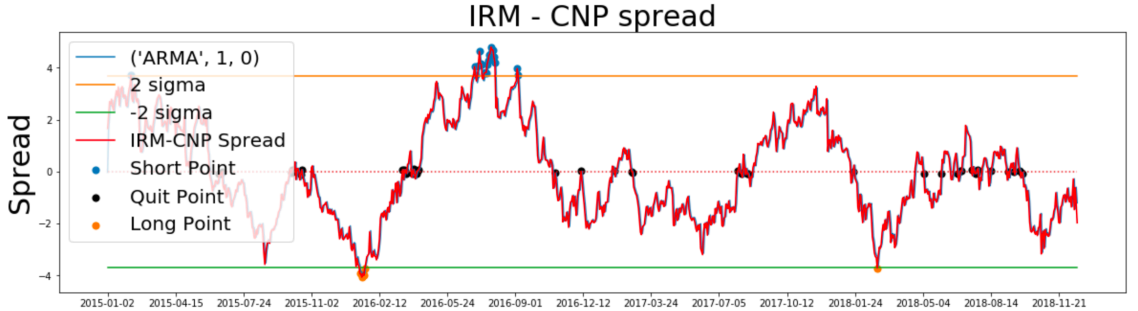
Appendix:

Performance of Cointegration method in Part C:

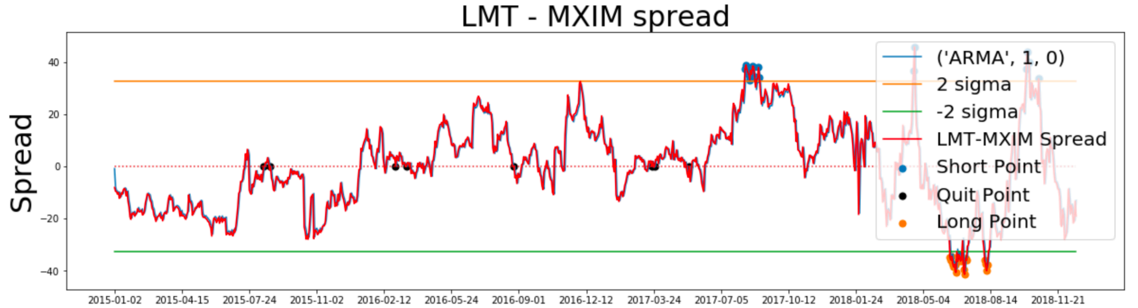
KR - URI spread



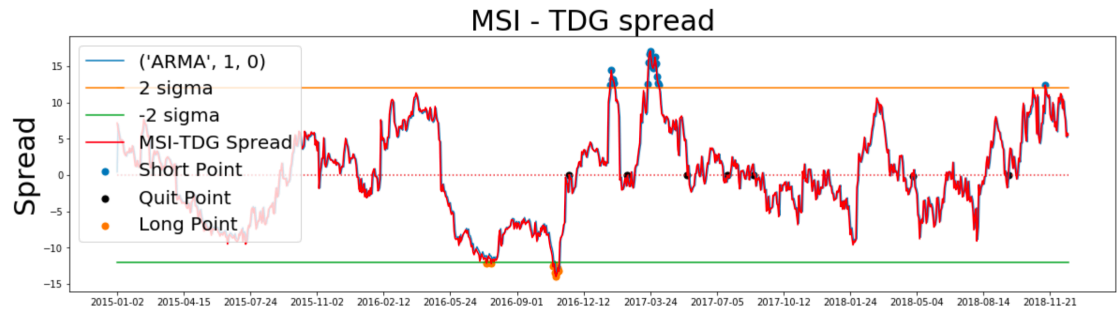
IRM - CNP spread

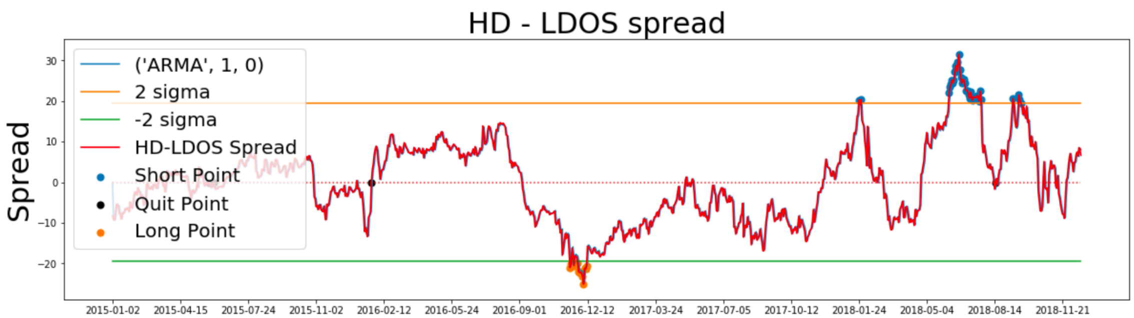


LMT - MXIM spread

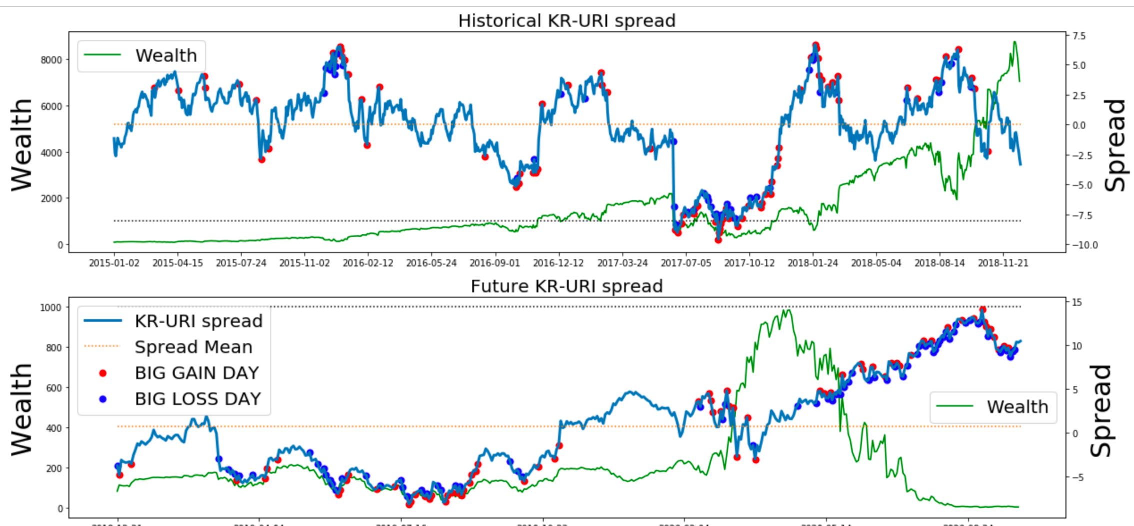


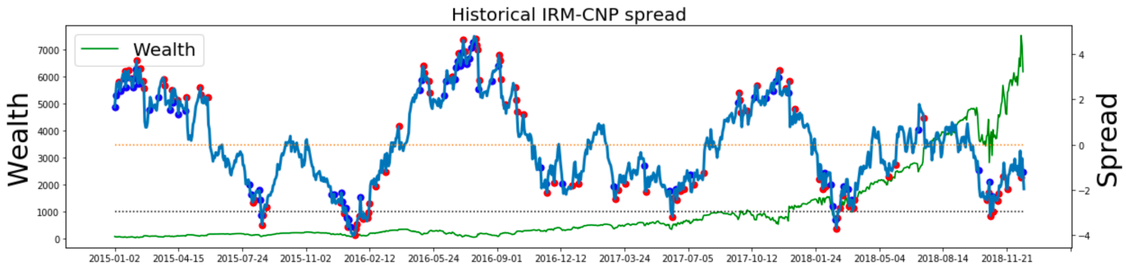
PCAR - ALL spread

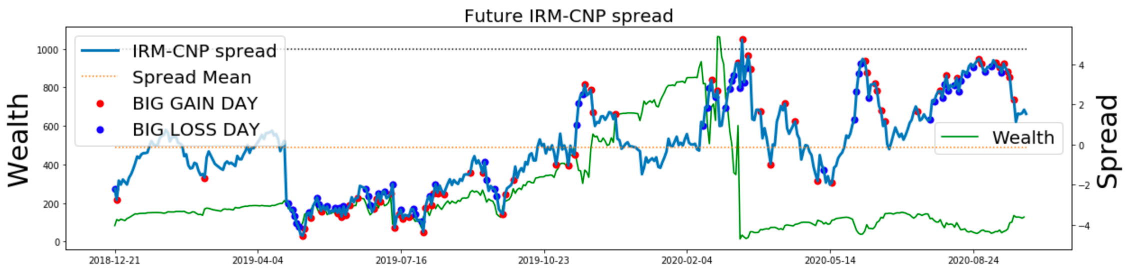


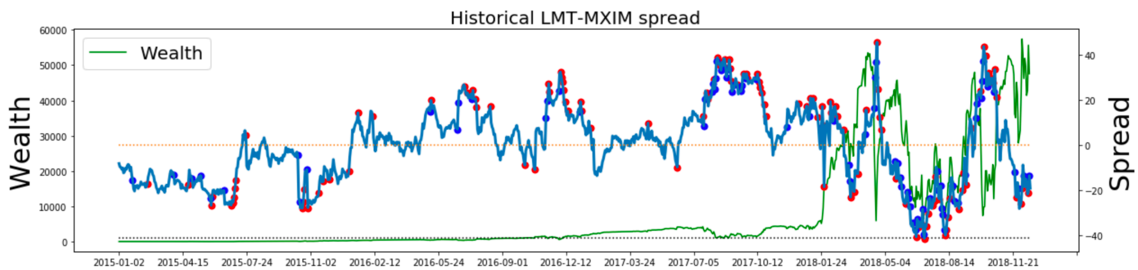


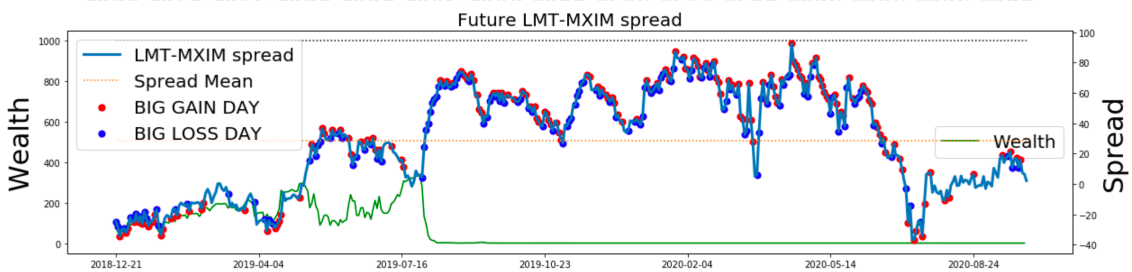
<u>Performance of Optimal Control method in Part D:</u>

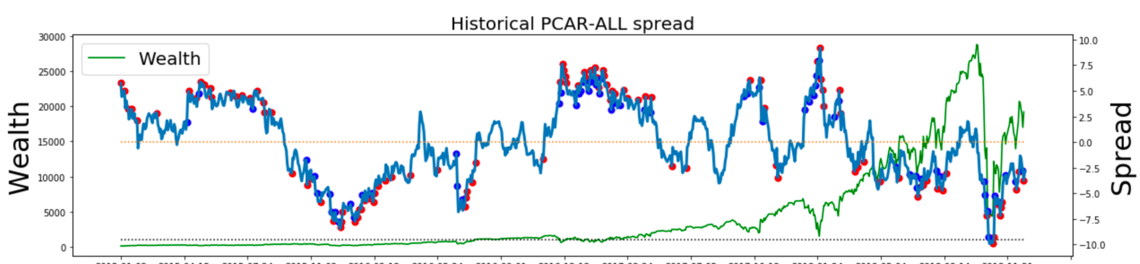


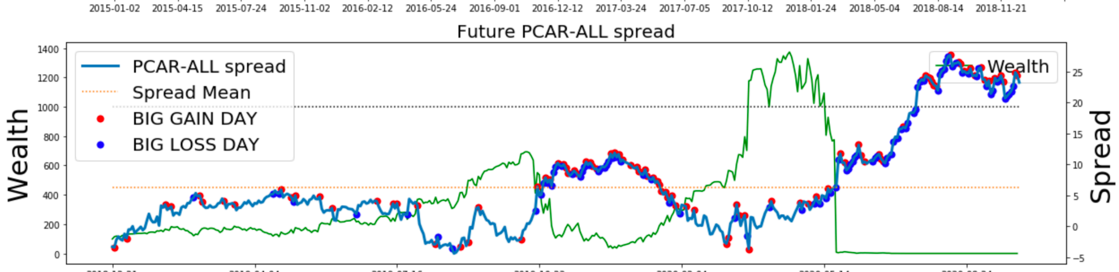

Historical IRM-CNP spread
Wealth
Wealth
Spread

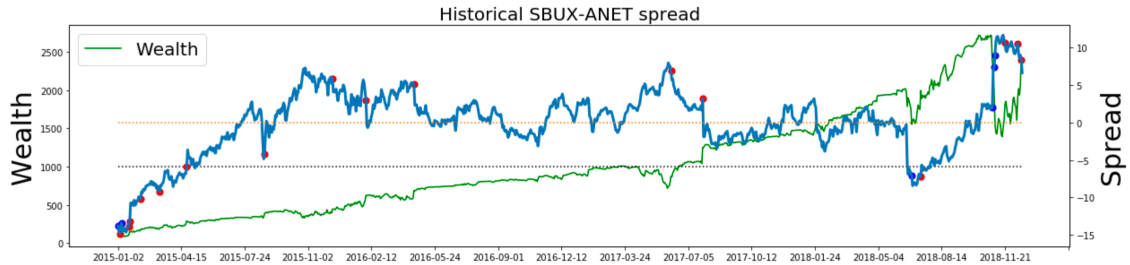

Future IRM-CNP spread
IRM-CNP spread
Spread Mean
BIG GAIN DAY
BIG LOSS DAY
Wealth
Wealth
Spread

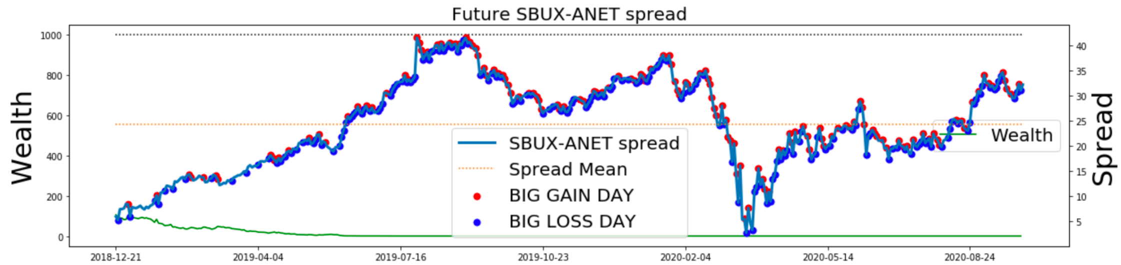

Historical LMT-MXIM spread
Wealth
Wealth
Spread

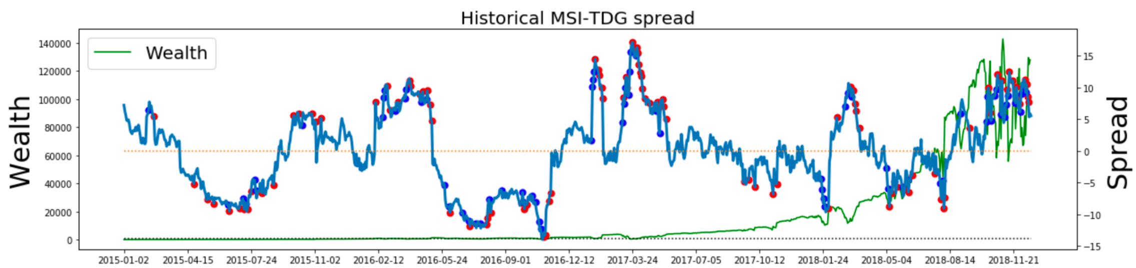

Future LMT-MXIM spread
LMT-MXIM spread
Spread Mean
BIG GAIN DAY
BIG LOSS DAY
Wealth
Wealth
Spread

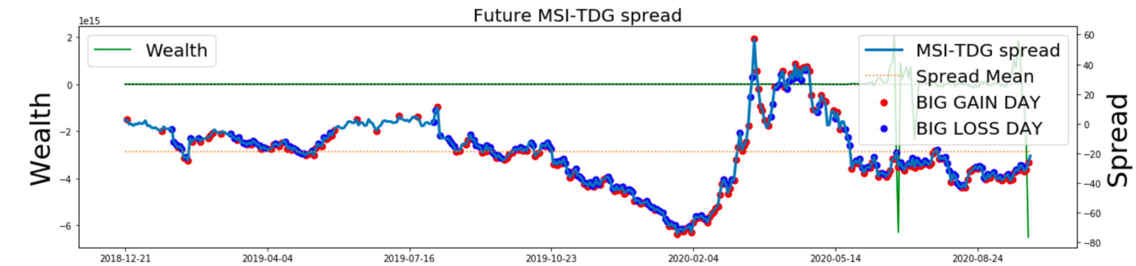

Historical PCAR-ALL spread
Wealth
Wealth
Spread


Future PCAR-ALL spread
PCAR-ALL spread
Spread Mean
BIG GAIN DAY
BIG LOSS DAY
Wealth
Wealth
Spread

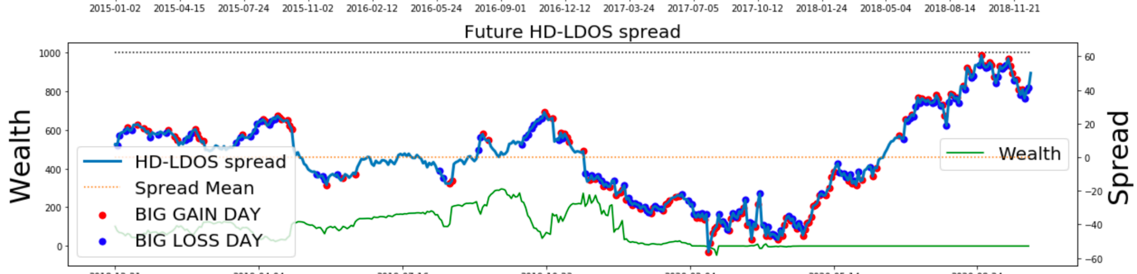

Historical SBUX-ANET spread
Wealth
Wealth
Spread


Future SBUX-ANET spread
SBUX-ANET spread
Spread Mean
BIG GAIN DAY
BIG LOSS DAY
Wealth
Wealth
Spread


Historical MSI-TDG spread
Wealth
Wealth
Spread


Future MSI-TDG spread
Wealth
MSI-TDG spread
Spread Mean
BIG GAIN DAY
BIG LOSS DAY
Wealth
Spread


Historical HD-LDOS spread
Wealth
Wealth
Spread


Future HD-LDOS spread
HD-LDOS spread
Spread Mean
BIG GAIN DAY
BIG LOSS DAY
Wealth
Wealth
Spread