CQF Final Project: Long/Short Trading Strategy Design & Backtest

Alexys Pereira

Introduction

This project is focused on implementing a Long/Short trading strategy designed using statistical analysis. This trading strategy consists on evaluating if a pair of assets that economically share similarities has constant spread i.e. if the difference between the prices of these two assets have constant mean and variance through time.
This price difference is named as spread, and we can evaluate if this difference has constant first and second moments through cointegration analysis.
Three pairs of assets will be studied:

1. **Royal Bank of Canada (RBC Bank) and Toronto-Dominion Bank (TD)**: these are the two largest Canadian banks and both are classified as G-SIBs, furthermore due to the market concentration in Canada's banking system, it is intuitive that these two assets have constant spread due to the similar nature of their business i.e. both are large and ultimately one can replace the other from an investor perspective.
2. **Japanese Yen (JPYUSD) and Gold Future (GC)**: there is a known relationship between those two assets that they are usually considered as safe-havens in periods of crisis due to stability of Japanese economy and Gold being a physical assets
3. **Canadian Dollar (CADUSD) and Oil Future**: Canada is one of the largest producers of oil in world and its currency has been observed to follow Oil prices fairly closely as oil is a large component of the country's economy, hence by evaluating if cointegration exists we can exploit any potential deviation for profits

A large component of the project is the implementation of numerical procedures from first principles, the list below shows the procedures implemented by the autor:

- **Ordinary Least Squares Regression (OLS) in Matrix Format**: coefficient estimation, standard error and p-value calculation, AIC and BIC calculations.
- **Augmented-Dickey Fuller Test**: equation specification, optimal lag selection and reject/accept decision - Mackinnon T-statistic and p-value calculation were used from statsmodels package since it was not covered in the course
- **Error Correction Model**
- **Ornstein-Uhlenbeck Process fitting**
- **Engle-Granger cointegration procedure**
- **Backtesting and risk assessment framework**:
  - Signal Generation
  - P&L Computation
  - Rolling Alpha, Beta and Sharpe Ratios

- Drawdown, VaR and ES calculation

Methodology

Cointegration parameter estimation: Engle-Granger Procedure

The methodology adopted for this Long/Short trading strategy design in based on pairs cointegration procedure Engle-Granger procedure to determine if two series are cointegrated.
Two series are said to be cointegrated if a linear combination between them is stationary, mathematically this can be expressed as:

$$Z_t = Y_t - \hat{\beta} X_t - c$$

If Z_t is stationary then X and Y are said to be cointegrated. Note that variable Z in fact is the residual from the Ordinary Least Square Regression of Y on X.

The Engle-Granger procedures has three main steps:
1. Estimate a linear regression of Y on X via OLS to extract the linear combination between the two series, which is given by the vector $[1, -\hat{\beta}]$
2. Test if the residual Z is a stationary using the Augmented Dickey Fuller Test (ADF) without trend term
The ADF test consists in determining if the given series Z has a unit-root i.e. if it is mean-reverting, this means that the level of Z_t can predict the change in ΔZ_t therefore mean-reversion exists.
Mathematically we define the test as an autoregression via OLS of ΔZ_t on its lagged level and differences i.e.

$$\Delta Z_t = (\rho - 1) Z_{t-1} + \sum_{k=0}^{k=K} \phi_k \Delta Z_{t-k} + \alpha + \epsilon_t$$

$$H_0 : \rho - 1 = 0 \; has \; Unit \; Root$$
$$H_1 : \rho - 1 \neq 0$$

The number of lags K is defined by fitting several equations and the one with the lowest AIC is selected as the best equation in which a decision will be made if the series is stationary or not.

3. Fit an Error Correction Model via Linear Regression with the following specification:

$$\Delta Y_t = \phi \Delta X_t - (1 - \alpha) Z_{t-1}$$

The Error Correction portion is the lagged residual Z_{t-1} and its coefficient $(1 - \alpha)$ must the statistically significant

If residual Z is stationary (2) and the error correction coefficient is statistically significant (3), hence we can say that series X and Y are cointegrated and that there is an underlying factor between them that makes their spread revert to its mean

The Engle-Granger procedure is compute in both direction i.e. $Y_t = \hat{\beta} X_t + c$ and $X_t = \hat{\beta} Y_t + c$
The direction that is selected to produce the trades is the one that has the highest absolute T-Statistic values in the ADF test

Assessment of quality of mean reversion: Ornstein-Uhlenbeck process

After confirming that two securities are indeed cointegrated i.e. there is mean-reverting difference between their prices (spread), we can design trades based on this mean reversion by estimating the parameters of the Ornstein-Uhlenbeck (OU) stochastic process.
Estimating this process provides two main insights:
a) The speed of the mean-reversion or its half-life: measurement in days of how long does it take for the spread/residual to return to its long-term mean
b) Entry and exit points of the position by measuring the diffusion (standard deviation) of the OU process: with this estimate we can determine what is the proper time to enter a position.

The OU process consisting in solving the following Stochastic Differential Equation (SDE):

$$de_t = -\theta(e_t - \mu_e)dt + \sigma_{OU} dX_t$$

SDE Solution:

$$e_{t+\tau} = (1 - e^{-\theta\tau})\mu_e + e^{\theta\tau} e_t + \epsilon_{t,\tau}$$

But we are interested in the parameters θ and μ_e because they refer to mean-reversion speed and equilibrium spread
To extract these parameters we can simply fit an AR(1) via OLS on the spread estimated via Cointegration, mathematically we compute the spread as

$$e_t = Y_t - \hat{\beta} X_t - \epsilon_t$$

And then using the output of this regression, we fit the following AR(1) process:

$$e_t = Be_{t-1} + C$$

Hence,

$$B = e^{-\theta\tau} \Rightarrow \theta = -\frac{\ln B}{\tau}$$

$$C = (1 - e^{-\theta\tau})\mu_e \Rightarrow \mu_e = \frac{C}{1 - B}$$

Finally the last parameters are the standard error of the equation that will be used for determining the upper and lower bounds

$$\sigma_{eq} = \sqrt{\frac{SSE\ \tau}{1 - e^{-2\theta\tau}}}$$

And also the Half-life, which measures on average how long does it take for the spread to return to its equilibrium level μ_e

$$\text{Half-life} = \frac{\ln 2}{\theta}$$

Trade Design

The trade strategy consisting in going long and short in each of series that are part of the cointegrated pair. There are three main decisions that need to be taken before doing the trade

1. Which security to short and which to go long?
2. How much of my capital my capital should I invest in each asset?
3. When should I enter and exit the trade?

To answer each of these questions we need to first define the upper/lower trading bounds:

$$Upper\ Bound = \mu_e + Z\sigma_{eq}$$

$$LowerBound = \mu_e - Z\sigma_{eq}$$

The bounds are symmetrical around the equilibrium point and are $Z\sigma_{eq}$ distant from it. This means that if the spread is greater than the Upper Bound or lower than the Lower Bound we start a trade, hence a signal is generated.
If the spread is greater than the Upper Bound, this means that there is a tendency for the spread to return back to its equilibrium, hence we should short asset Y (the dependent variable on the regression) and go long $\hat{\beta}$ units of asset X.
But if spread is lower than the Lower Bound then there is a tendency for the spread to increase i.e $Y - \hat{\beta}X$ is increasing there fore we go long on Y and short $\hat{\beta}$ X.
Mathematically:

$$If\ Spread_t > Upper\ Bound = \mu_e + Z\sigma_{eq} \Rightarrow Weights = [-100\%Y, \%\hat{\beta}X]$$

$$If\ Spread_t < Lower\ Bound = \mu_e - Z\sigma_{eq} \Rightarrow Weights = [100\%Y, -\%\hat{\beta}X]$$

$$Exit\ Position\ when\ Spread_t \approx \mu_e$$

Finally we exit our position when the spread is close to the equilibrium level μ_e
Because there is a hyperparameter Z involved in this trade design, the optimal threshold level will be determined by testing different levels of Z and the one selected will be the one the provides the best out-

of-sample performance in terms of Alpha and Sharpe Ratio. In total, 7 values for Z will be explored [0.2, 0.3, 0.4, 0.5, 0.6, 0.7] - we want tight spreads to generate a higher number of trades over a period of time, otherwise there will be few trades and as time passes the estimated equilibrium might shift.

Performance Measurement

To assess the optimal Z level to start trades I will several risk-return and P&L metrics, both for the whole period as well as their six-month rolling measurement. The reason for that being that by computing rolling metrics we can assess how volatile the strategy is through time, theoretically they should be fairly stable because long/short strategy such as pairs trading are not directional.
The following metrics will be used to assess strategy performance:

- Cumulative and Annualized Return
- Cumulative and Annualized Volatility
- Annualized Alpha and Beta
- Six-month rolling and average Sharpe Ratio
- Six-month rolling and average Alpha
- Six-month rolling and average Beta
- Maximum Drawdown
- Six-month rolling and average 10-day VaR 99%

Data preprocessing

The preprocessing of the time-series will be minimal limited to:

- Filling null values with previous value (forward filling)
- Price normalization such that the prices such that all start at one. This manipulation is useful because addresses difference in scales between the securities and has the intuition that if I were to invest one dollar in each security how much I would earn, hence because the trading weights will be based on the coefficients then we can apply directly to build the portfolio

Data and Backtesting/Validation Strategy

The data utilized is daily market prices extracted from Yahoo Finance using the pandas_datareader package between Jan 1st, 2010 and Feb 1st, 2020. This period has been selected as it was not affected by the Great Financial Crisis, COVID-19 Pandemic nor Russia-Ukraine war, this is important because cointegration are long-term trends and these events breakdown these patterns since the trade strategy assumes that these relationship will be stable moving forward. In other words, the mean-reversion properties change because correlation between assets increase such that it does not mean-reverts and the equilibrium μ_e itself changes.
The second points refers to backtesting/validation strategy: I have decided to follow to divide the data into three smaller data sets

- **Training Set - Jan 1st, 2010 to Dec 31st, 2016:** this period will be used to estimate all the relevant parameters that will be used to design trades in the following periods

- **Validation Set - Feb 1st, 2017 to Feb 1st, 2019:** the validation set serves to evaluate the model out-of-sample performance to select the optimal strategy (optimal Z) such that maximizes Alpha and Sharpe Ratio
- **Test set - Mar 1st, 2019 to Feb 1st, 2020:** the gold standard where the results are in fact report if the strategy designed behaves as expected (or not and why).

Note that between each data set there is a one-month gap to minimize data leakage in order to maintain fairness.

Implementation and Results

With the data extracted and methodology laid out, we can proceed to the implementation and results cointegration for pairs tradings.
We start by depicting in **Table 1** a small sample of the data set with prices normalized such that the initial point equals 1.

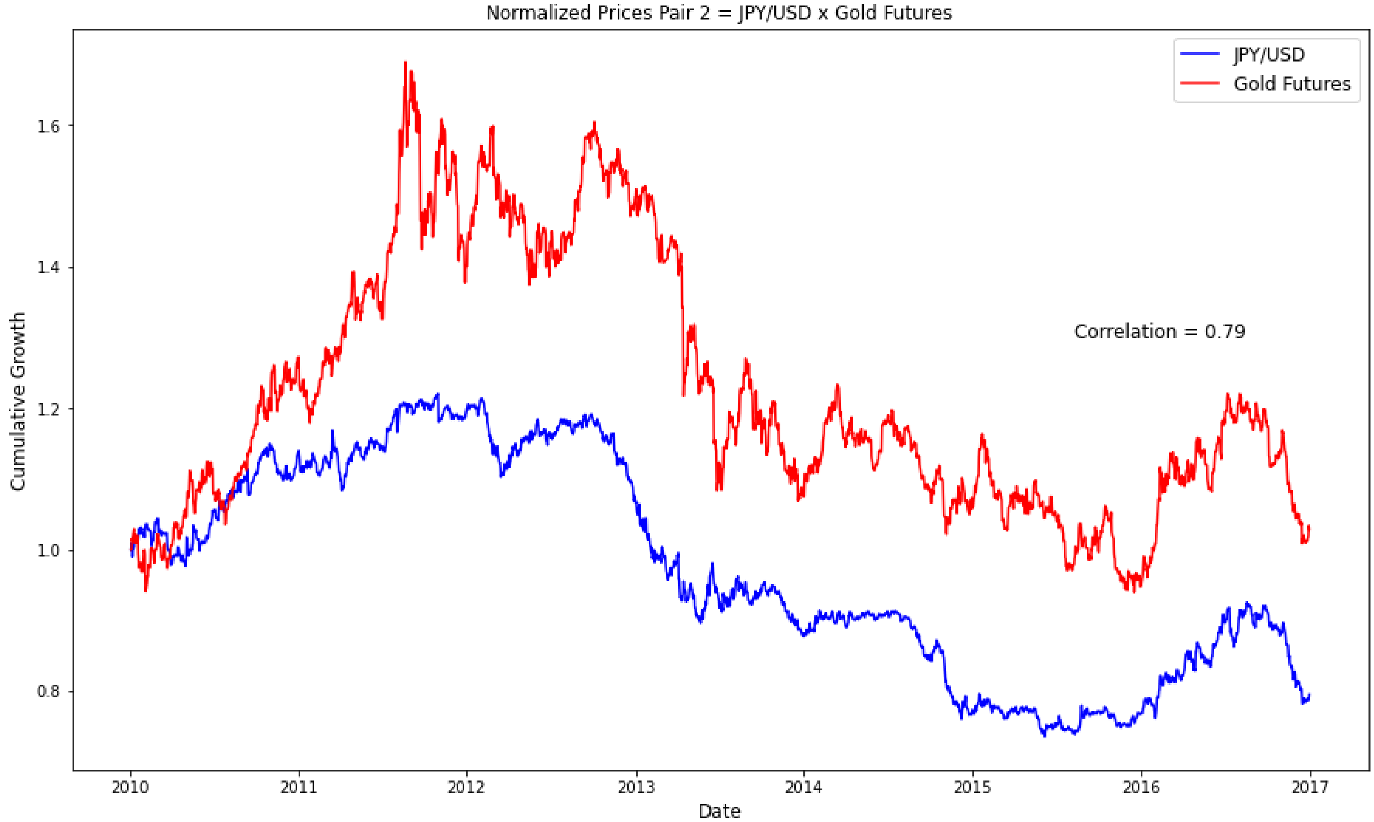
Table 1: Sample Data Used

Symbols	Date	CADUSD=X	BZ=F	RY.TO	TD.TO	JPYUSD=X	GC=F
1	2010-01-04 00:00:00	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
2	2010-01-05 00:00:00	1.001443	1.005866	0.984634	0.990402	1.009936	1.000358
3	2010-01-06 00:00:00	1.008819	1.022092	0.977393	0.980652	1.001505	1.016284
4	2010-01-07 00:00:00	1.006770	1.017349	0.978276	0.972120	0.990417	1.013778
5	2010-01-08 00:00:00	1.010876	1.015602	0.976510	0.970902	0.998219	1.018341

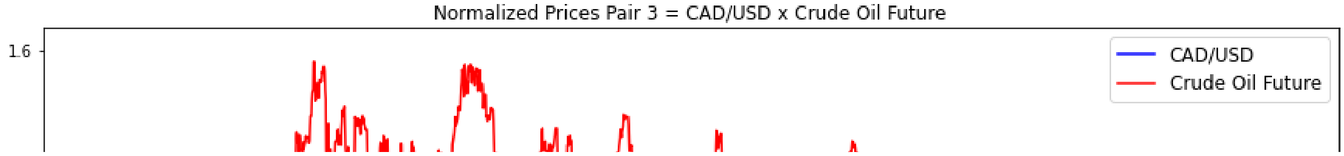
The charts below plots the price evolution of each trading pair (three in total) for the training period (Jan, 2010 - Dec, 2016). And we can noticed that all pairs exhibit a very high degree of correlation between the assets - RBC x TD Bank are 99% correlated, Japanese Yen x Gold 80% and Canadian Dollar x Oil 90%. However correlation does not guarantee cointegration of the series i.e. mean-reverting spread.

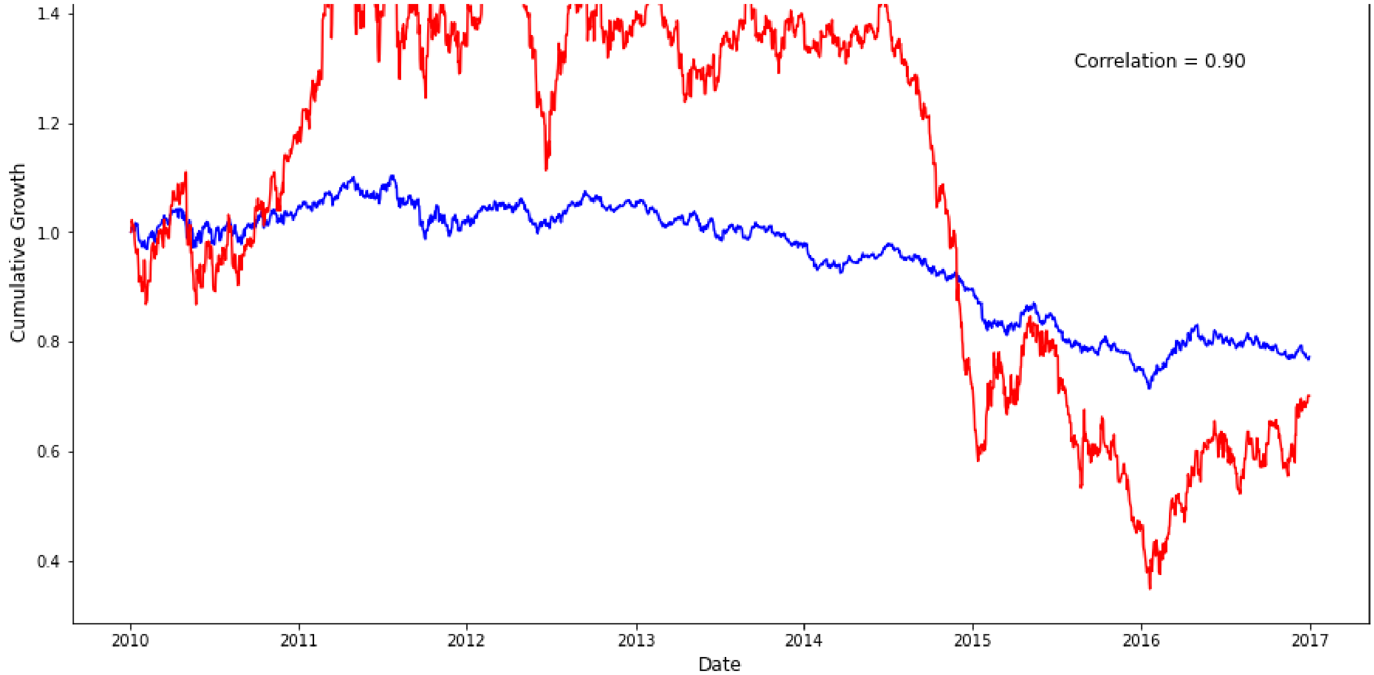
Normalized Prices Pair 1 = RBC x TD Bank

Normalized Prices Pair 2 = JPY/USD x Gold Futures



Normalized Prices Pair 3 = CAD/USD x Crude Oil Future





Pair 1: RBC Bank (RY.TO) and TD Bank (TD.TO)

The first pair we examine is between the two largest Canadian banks RBC and TD, and as depicted earlier their prices evolve at the same rate as shown by their almost perfect correlation.

Engle Granger Fitting

We started by applying the Engle-Granger procedure to evaluate if cointegration exists between RBC and TD. The first results is shown on **Table 2** that regress RBC on TD as well as TD on RBC
From **Table 2** we see that both regressions have highly significant coefficients.

Table 2: RY.TO and TD.TO OLS Regression Results

	Variable				RY.TO	Variable				TD.TO
Statistic		**Beta**	**P-Value**	**SE**	**T-Stat**		**Beta**	**P-Value**	**SE**	**T-Stat**
0	Constant	0.021	0.000	0.005	4.067	Constant	0.021	0.001	0.006	3.395
1	TD.TO	0.827	0.000	0.003	262.639	RY.TO	1.175	0.000	0.004	262.639

Table 3 is the ADF test for each equation, and notice that the spreads for both equations are stationary series due to their high T-Statistic. Equation 1 has a T-Statistic of -3.62 whereas equation 2 T-statistic is -3.66, both significant at 1% confidence level.
Because Equation 2 ($TD = f(RBC)$) has a larger absolute T-statistic I decide to move forward with this equation. Hence we move forward Equation 2 for the next steps

Table 3: RY.TO and TD.TO ADF Test Results

	Equation 1:	RY.TO = TD.TO + Constant		Equation 2:	TD.TO = RY.TO + Constant	
Null Hypothesis	Series has unit root			Series has unit root		
T-Statistic	-3.632942			-3.669791		
P-Value (MacKinnon)	0.005162			0.004559		
Optimal Lag	0			0		
Confidence Level	1%	5%	10%	1%	5%	10%
Mackinnon Critical Value	-3.433569	-2.862962	-2.567527	-3.433569	-2.862962	-2.567527
Reject/Not Reject H0	Reject	Reject	Reject	Reject	Reject	Reject
Stationary/Non Stationary	Stationary	Stationary	Stationary	Stationary	Stationary	Stationary

Table 4 exhibits the results for the error equation 2, dY refers to TD's first difference, dX is RBC first difference and Residual(-1) is the spread, mathematically this is the formulation

$$\Delta TD_t = \Delta RBC_t - (1-\alpha)(TD - \hat{\beta}RBC - c)_{t-1}$$

We can see from the table that Residual(-1) p-value is much lower than 1% hence there is evidence of change in prices due to past difference between TD and RBC
Therefore the final cointegration equation that will be used to design the trades is as follows:

$$Spread_t = TD_t - 1.175RBC - 0.021$$

Table 4: Equation 2: TD.TO = RY.TO + Const Error Correction Model Results

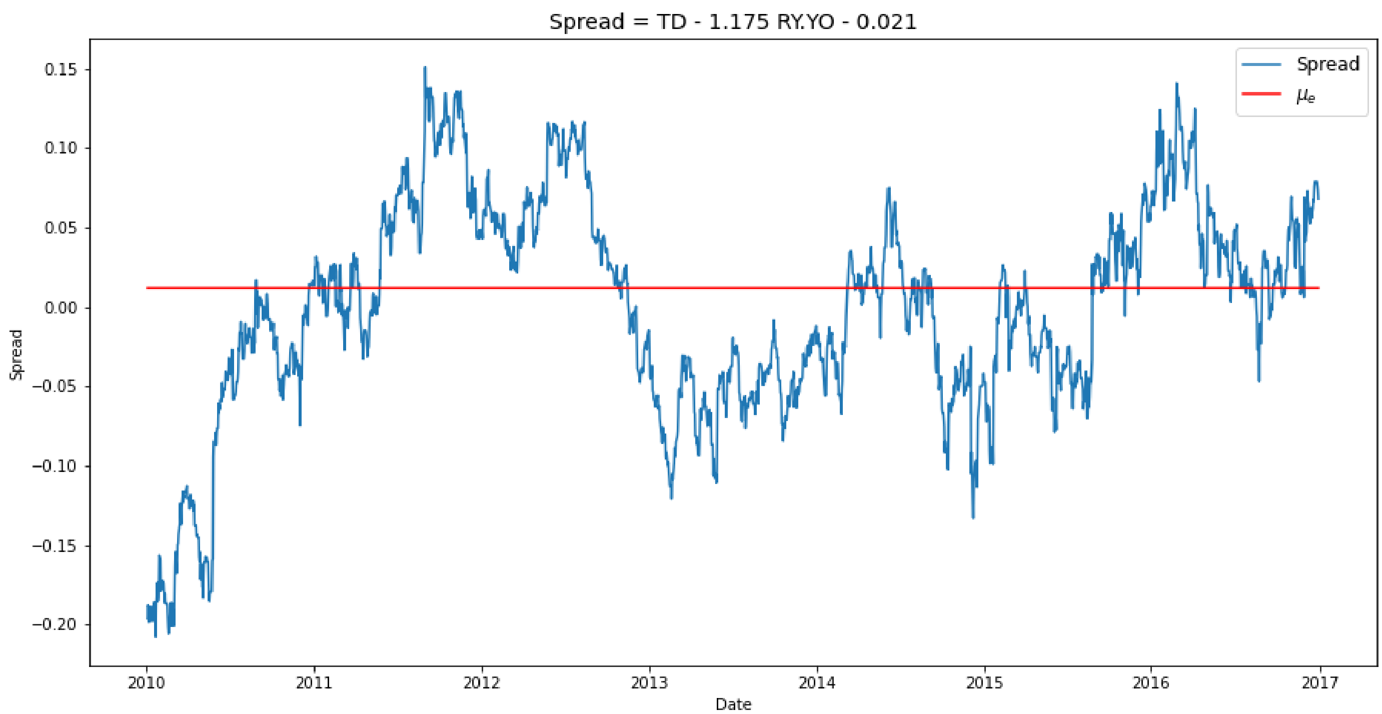
				dY
Statistic	**Beta**	**P-Value**	**SE**	**T-Stat**
Variable				
dX	0.889	0.000	0.015	59.996
Residual(-1)	-0.008	0.002	0.003	-3.090

Given that the evaluated series has passed all the required statistical tests, we can proceed to fit the OU process to estimate the trading parameters to be utilized.
Table 5 showcases all the relevant parameters with equilibrium ean of 0.012476 and standard deviation of 0.18, more than 10 higher than the equilibrium level. Another important parameter is the Half-life

Table 5: RY.TO and TD.TO OU Process Fitted Parameters

	Value
OU Process Parameters	
μ_e	0.011867
Half-Life (days)	63.098626
σ	0.173066
θ	2.768255

The chart below plots the spread evolution for the training period in blue and the equilibrium point μ_e, notice that there is a cyclical pattern in which the spread returns to its equilibrium characteristic of a stationary series



Backtesting Pair Trading

With the cointegration weights properly estimated and its statistical significance verified, I move forward to backtest the performance of my trading strategy by assessing how profitable each strategy is as a function of Z variable to define the lower and upper bounds.

$$If\ Spread_t > Upper\ Bound = 0.012476 + 0.182Z \Rightarrow Weights = [-100\%TD, \%\hat{\beta}RBC]$$

$$If\ Spread_t < Lower\ Bound = 0.012476 - 0.182Z \Rightarrow Weights = [100\%TD, -\%\hat{\beta}RBC]$$

$$Exit\ Position\ when\ Spread_t \approx 0.012476$$

In order to guarantee that trades will close, trades will be closed if

$$Spread_t = [0.75\mu_e, 1.25\mu_e]$$

Table 6 summarizes P&L metrics as well as the number of trades and the number of days opens. Notice that for Z = 0.2, 20 trades were performed with an annualized return of 17% and on average each trade lasts for 77 days (9 days more than the OU half-lfe).
As we increase the Z-score to 0.3 the number of trades decline drastically, pointing out that Z=0.2 is capturing a lot of noise instead of true signal. Also, as Z increases the number of days that the trade is open increases.

Table 6: RY.TO and TD.TO P&L Results: Jan 1st, 2010 - Dec 31st, 2016

	Strategy ID	Z-Score	# Trades	Annualized Return (%)	Annualized Volatility (%)	Cumulative Return (%)	Avg # Days Open Position	Total Trading Days
0	1.00	0.20	17.00	7.49	11.44	57.17	115.59	1965.00
1	2.00	0.30	13.00	8.35	11.50	59.05	140.00	1820.00
2	3.00	0.40	7.00	10.64	11.88	61.42	214.00	1498.00
3	4.00	0.50	6.00	12.38	11.76	65.61	227.33	1364.00
4	5.00	0.60	5.00	13.06	11.87	62.51	251.20	1256.00
5	6.00	0.70	5.00	14.81	11.64	65.92	231.20	1156.00

Table 7 summarizes Risk-Adjusted metrics for in-sample period, all in percentage terms,and we can see similar pattern as the regular P&L metrics.
It is interesting to note that Beta across all strategies is fairly low, hovering around between -0.1 and 0.18 confirming the initial hypothesis that long/short strategy should be market neutral. All Z options generate positive alpha between 17.8% and 24%.

Table 7: RY.TO and TD.TO Risk-Return Metrics: Jan 1st, 2010 - Dec 31st, 2016

	Value (%)					
Strategy	**Z-Score = 0.2**	**Z-Score = 0.3**	**Z-Score = 0.4**	**Z-Score = 0.5**	**Z-Score = 0.6**	**Z-Score = 0.7**
Metric						
Annual Return	7.49	8.35	10.64	12.38	13.06	14.81
Annual Vol	11.44	11.50	11.88	11.76	11.87	11.64
Cumulative Return	57.17	59.05	61.42	65.61	62.51	65.92
Alpha (Annual)	7.83	8.75	10.87	12.35	12.16	14.16
Sharpe	38.26	45.14	61.25	75.42	79.60	93.93
Beta	7.86	9.08	9.30	11.95	12.84	8.34
Beta P-Value	0.02	0.00	0.01	0.00	0.00	0.31
Alpha P-Value	8.57	6.57	4.54	2.77	3.97	1.96
Max Drawdown	15.49	12.92	10.87	9.99	9.99	9.99
1-Day VaR 99%	1.65	1.65	1.70	1.67	1.69	1.65
10-Day VaR 99%	4.99	4.98	5.07	4.96	4.98	4.82
1-Day ES 99%	1.89	1.90	1.95	1.92	1.94	1.90
10-Day ES 99%	5.76	5.76	5.88	5.75	5.79	5.61

Table 8 summarizes P&L metrics for the validation/out-of-sample period, all in percentage terms. The most interesting point is that there is a degradation of performance for wider interval (larger Z values) due to larger average trade half-life. Pointing out that this strategy to be sustainable must be performed in higher frequencies to avoid being caught in change of regime or the relationship breaking down.

Table 8: RY.TO and TD.TO P&L Results: Feb 1st, 2017 - Feb 1st, 2019

	Strategy ID	Z-Score	# Trades	Annualized Return (%)	Annualized Volatility (%)	Cumulative Return (%)	Avg # Days Open Position	Total Trading Days
0	1.00	0.20	6.00	2.92	8.93	6.18	107.17	643.00
1	2.00	0.30	5.00	2.64	8.91	5.39	123.80	619.00
2	3.00	0.40	5.00	4.50	8.96	8.96	119.20	596.00
3	4.00	0.50	4.00	4.74	8.94	9.09	143.75	575.00
4	5.00	0.60	3.00	2.64	9.17	4.52	173.33	520.00
5	6.00	0.70	2.00	0.78	9.07	1.29	255.50	511.00

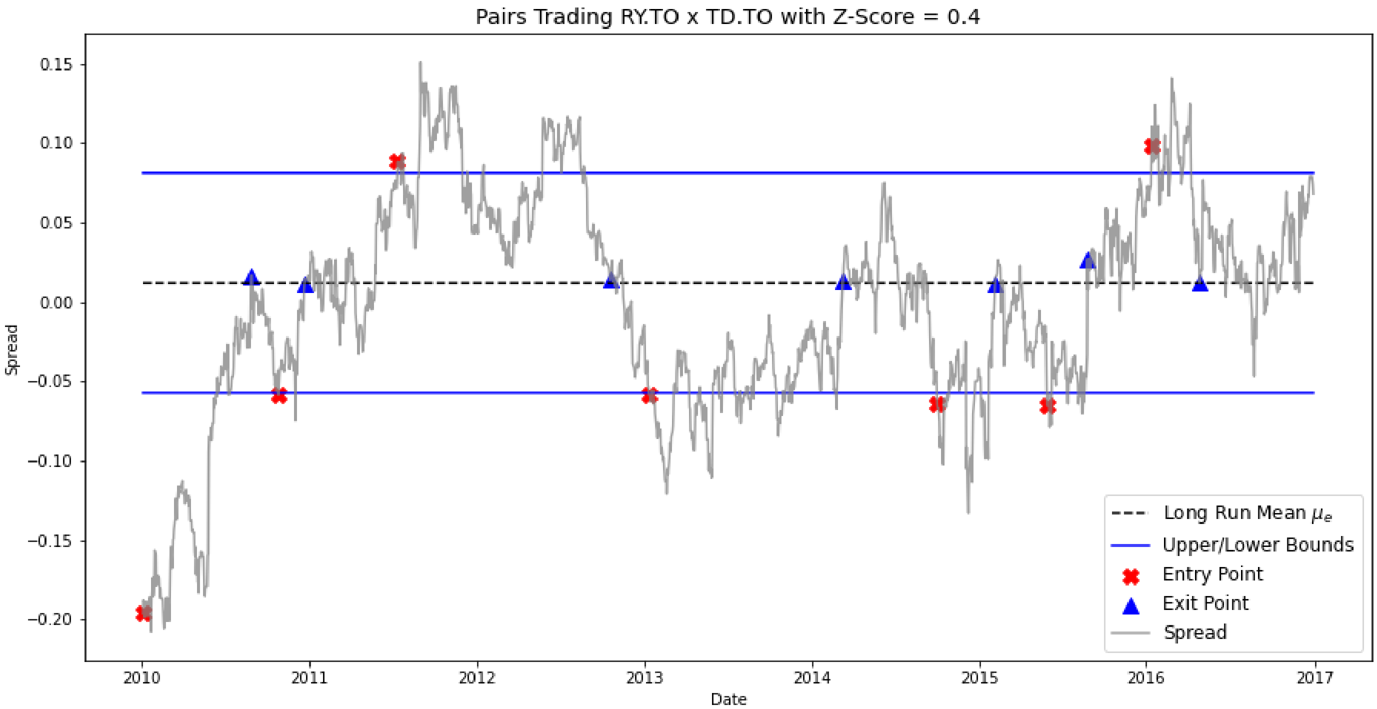
Table 9 summarizes the risk-adjusted metrics also for the validation period, and we can see that for lower Z values of 0.2, 0.3 and on the higher end of 0.6, 0.7 they have negative Sharpe ratios confirming that too tight spread capture noise rather than true signal, and to wide spreads make the strategy vulnerable to

breakdown in relationship. **Hence we choose Z = 0.4 as optimal level, since it provides the highest Alpha for the period of 3.43%.**

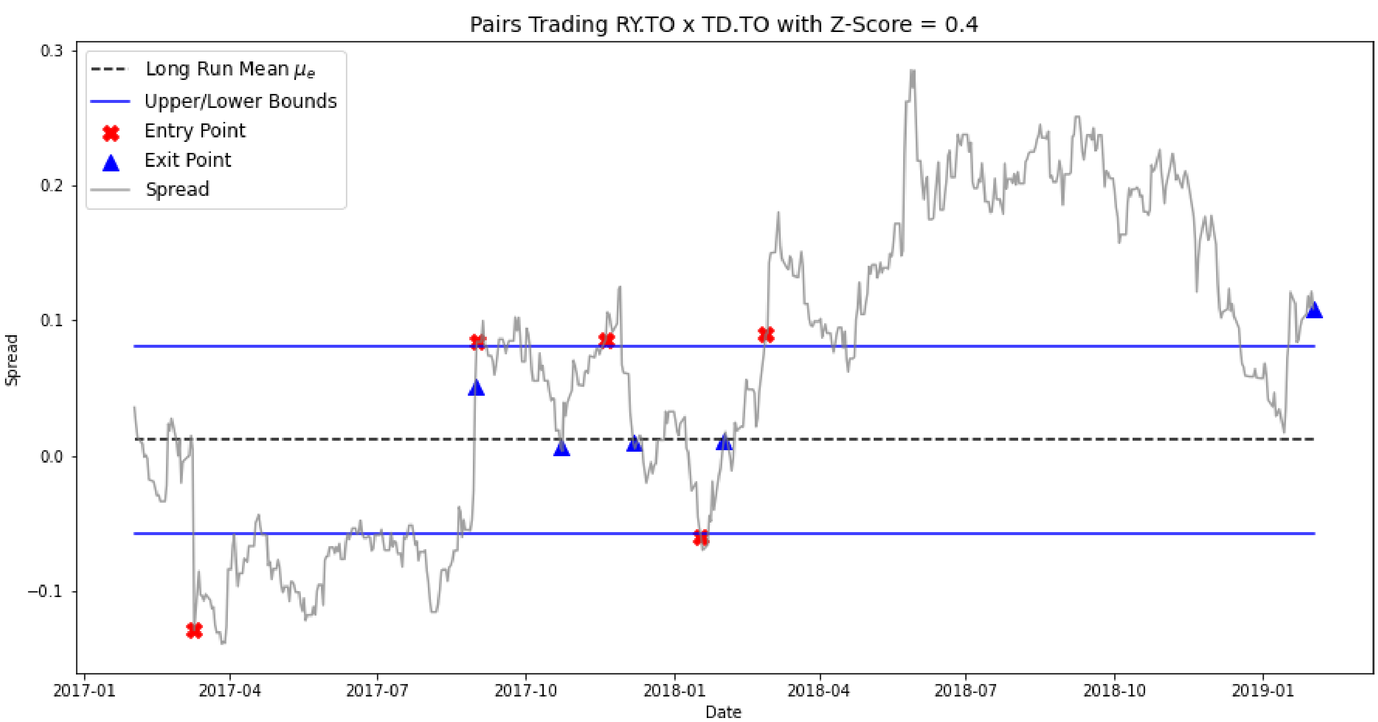
Table 9: RY.TO and TD.TO Risk-Return Metrics: Feb 1st, 2017 - Feb 1st, 2019

						Value (%)
Strategy	**Z-Score = 0.2**	**Z-Score = 0.3**	**Z-Score = 0.4**	**Z-Score = 0.5**	**Z-Score = 0.6**	**Z-Score = 0.7**
Metric						
Annual Return	2.92	2.64	4.50	4.74	2.64	0.78
Annual Vol	8.93	8.91	8.96	8.94	9.17	9.07
Cumulative Return	6.18	5.39	8.96	9.09	4.52	1.29
Alpha (Annual)	3.41	3.34	5.16	5.07	3.37	1.50
Sharpe	4.73	1.61	21.72	24.23	1.67	-18.65
Beta	10.93	11.78	11.85	12.82	12.81	12.19
Beta P-Value	1.22	0.71	0.74	0.42	0.62	0.88
Alpha P-Value	58.02	59.29	41.94	43.38	63.01	83.05
Max Drawdown	9.42	8.64	7.73	7.03	7.03	7.03
1-Day VaR 99%	1.30	1.29	1.29	1.29	1.33	1.32
10-Day VaR 99%	4.01	4.01	3.96	3.94	4.13	4.16
1-Day ES 99%	1.49	1.48	1.49	1.48	1.53	1.52
10-Day ES 99%	4.61	4.61	4.57	4.55	4.75	4.77

The chart below plots in grey the spread for the training (in-sample) period between 2010 and 2016, the blue lines are Upper/Lower bounds compute using the OU parameters with optimal Z=0.4. Notice that we can see that there is mean-reverting/stationary behaviour as the grey lines after breaching the blue are pulled towards the black dashed line which is the long run mean μ_e.
The trades exploit this behaviour, notice that the trades are initiated at the red crosses, whenever the grey line passes the blue ones, and the blue triangles are when the trades are closed. On interesting pint is that there are few trades which have very long lifespan and several shorter ones

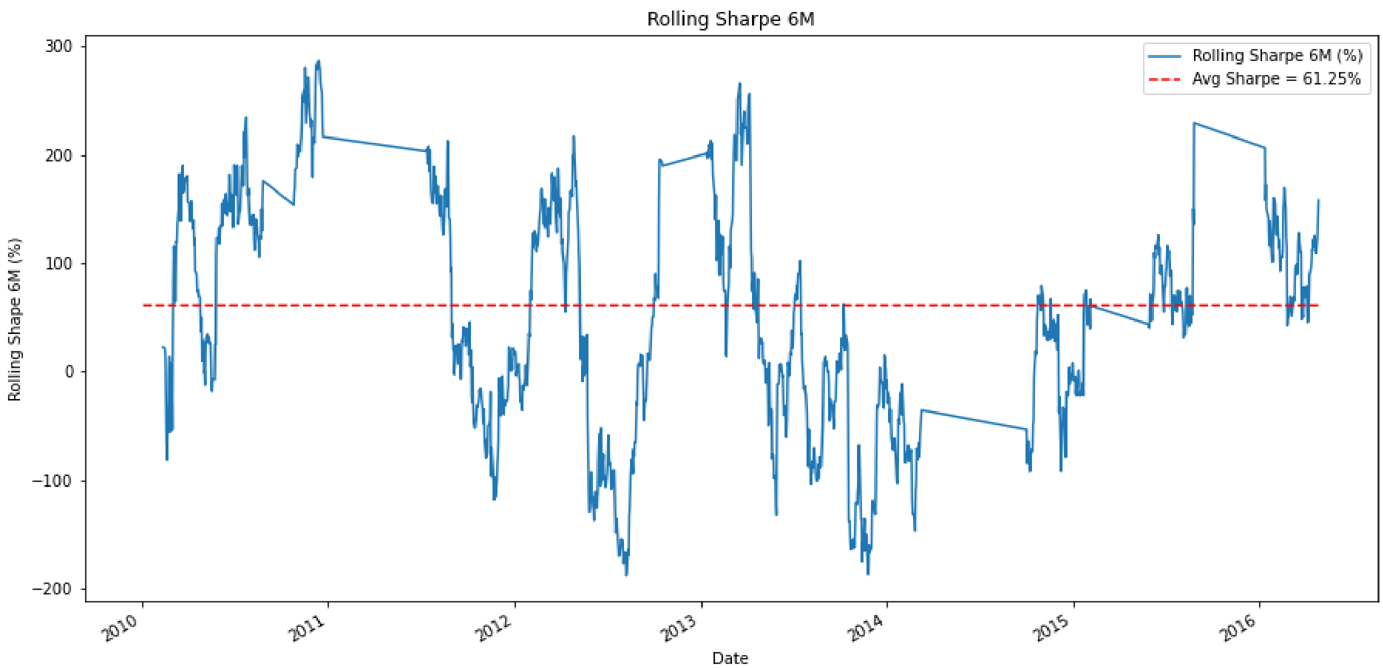


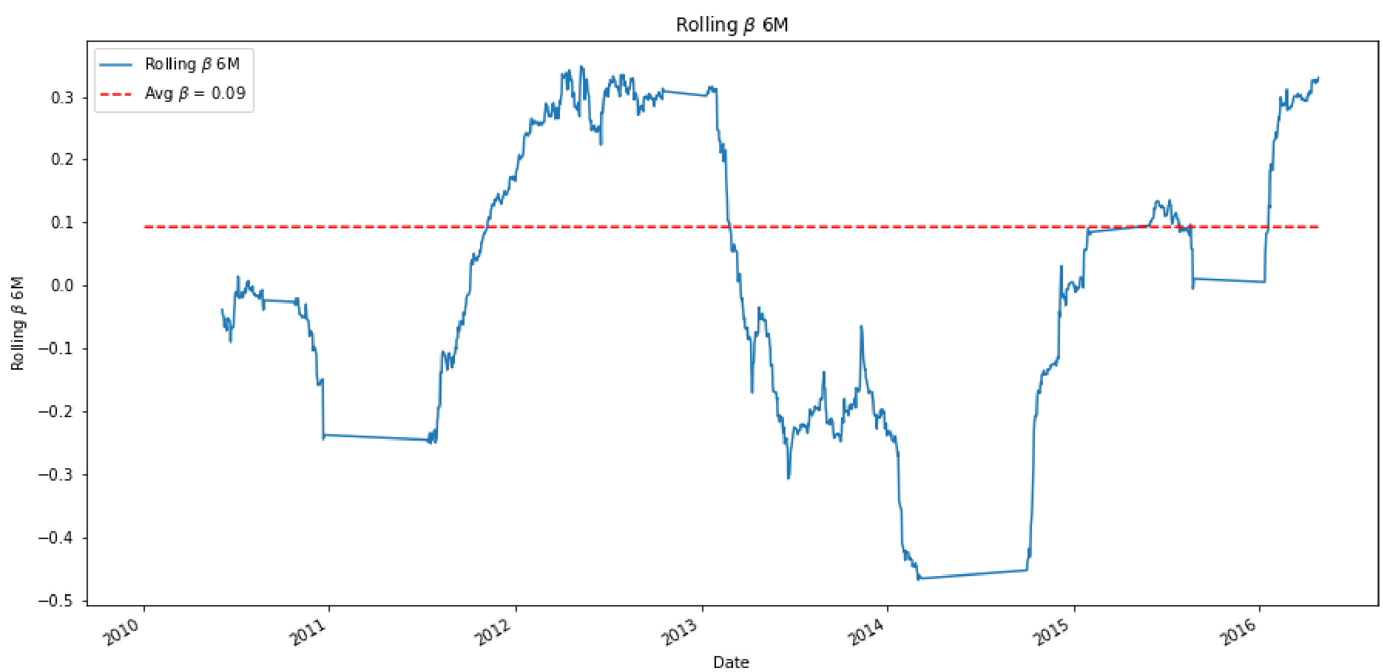
Now, we turn our attention for the validation period plotted on the chart below and we can see similar pattern as observed for the training period. Several short lived trades, and a few long lived ones that skew the average days a position is open.
However, after 2018 it seems that there is a breakdown in the cointegration as there is not mean-reversion for almost an year, this exhibits the necessity of restimating the equations using Kalman filter or doing the whole Engle-Granger procedure for recalibrating

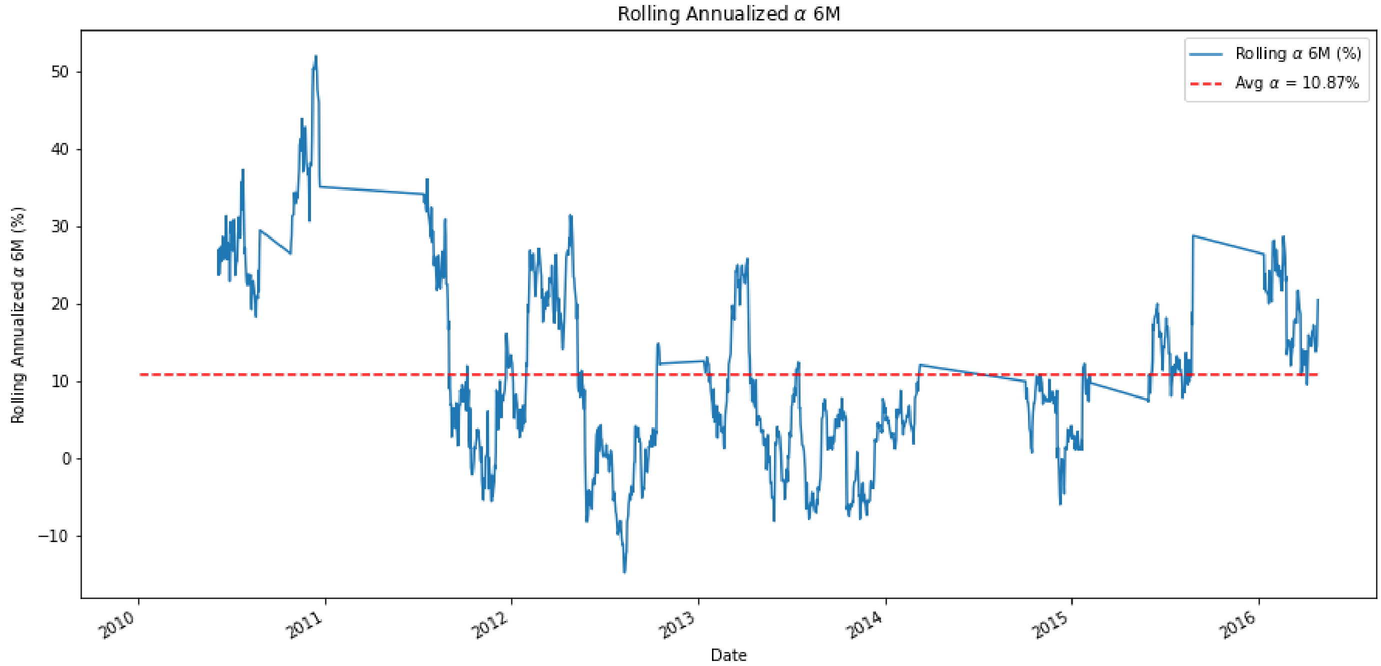


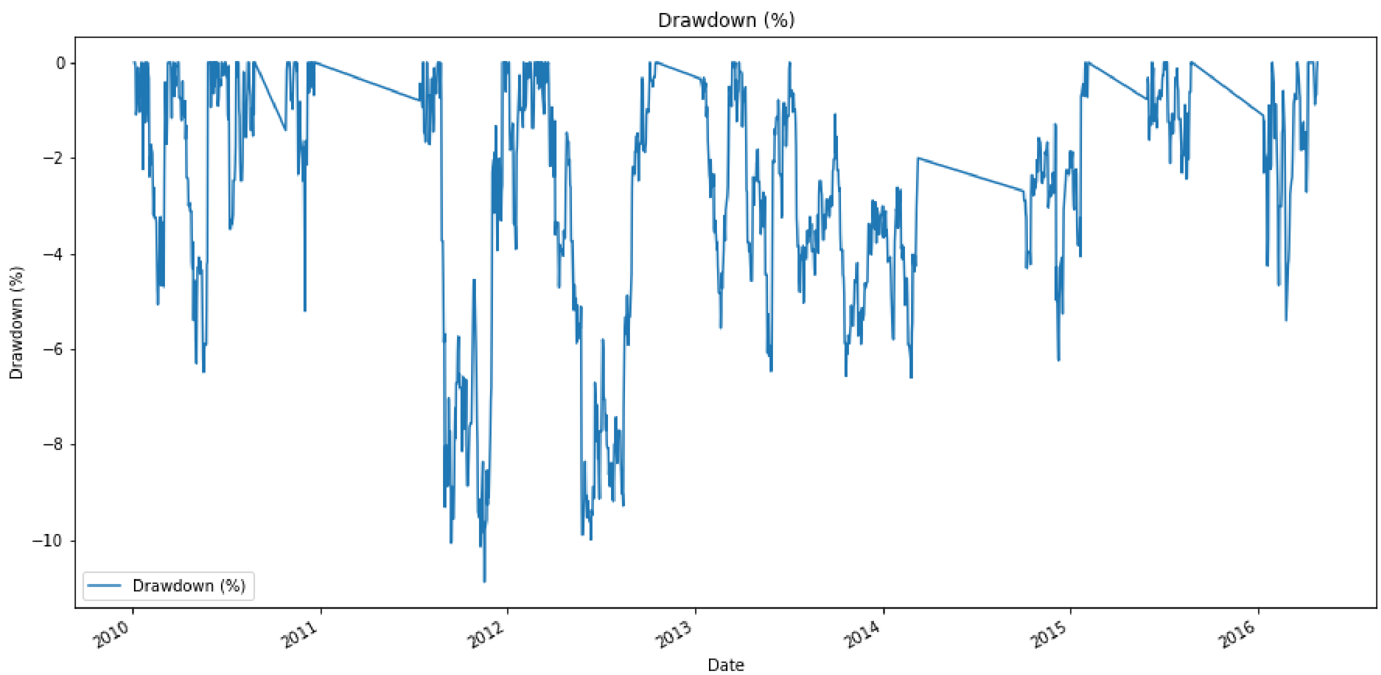
To complete the analysis we now move on to analyze rolling risk-adjusted metrics, the sequence of three plots below refers to the training period between 2010 and 2016.

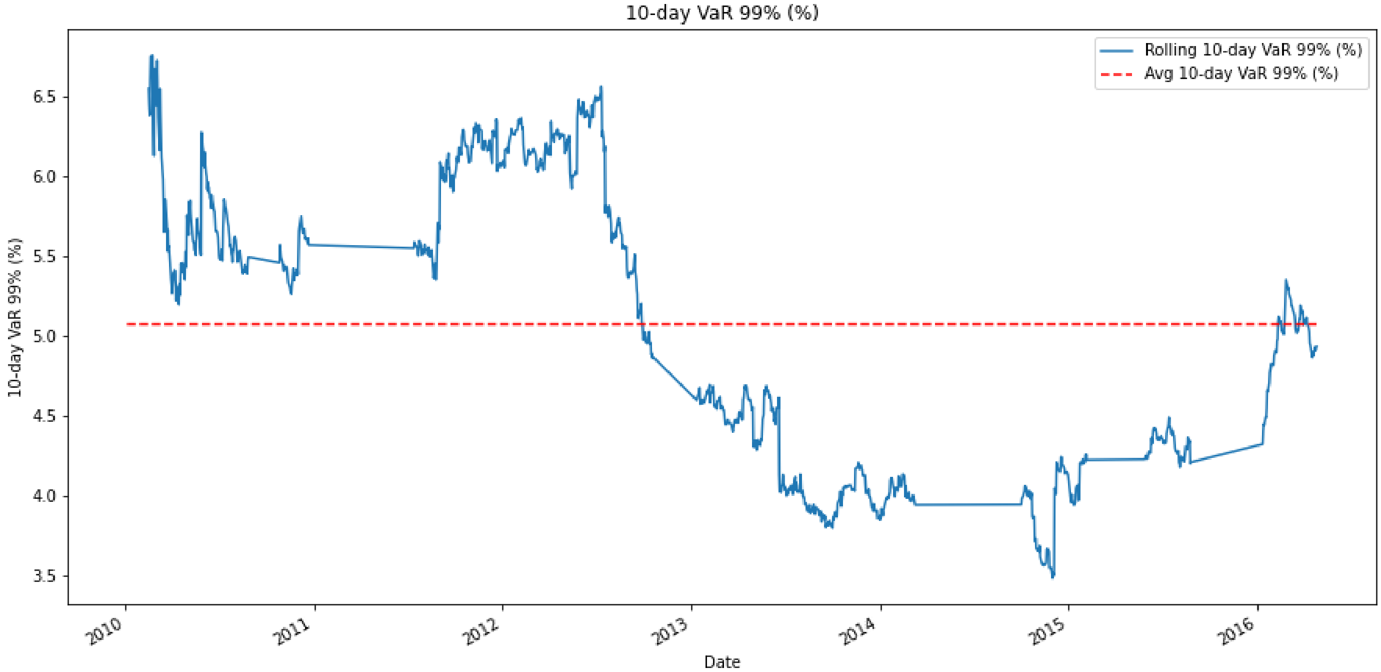
The first plot shows the rolling six-month Sharpe Ratio and it is clear that it exhibits a very high variability around the average of 61.25% which is not desirable and a function of trade that stay open for too long. The second and thirds plots are, respectively, the β and α in relation to the Toronto Stock Exchange Index (TSX) and similar pattern can be observed, a highly oscillating coefficients around the average.



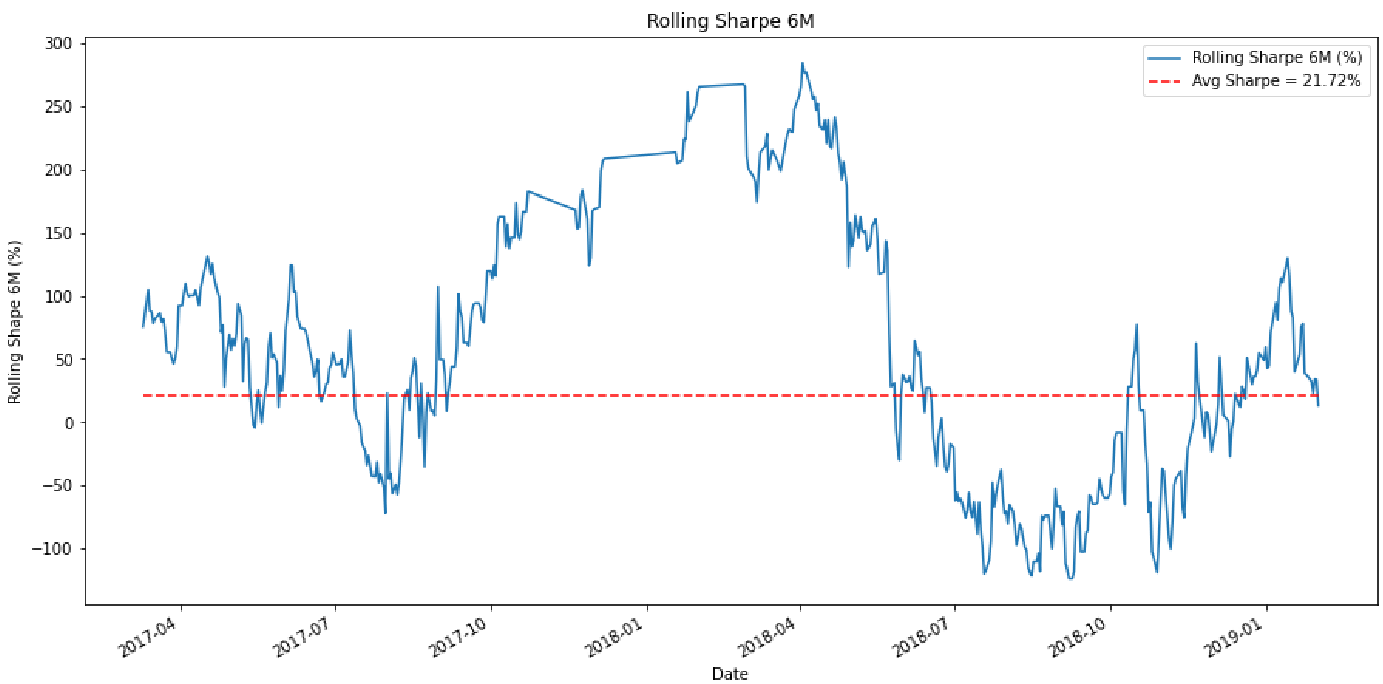


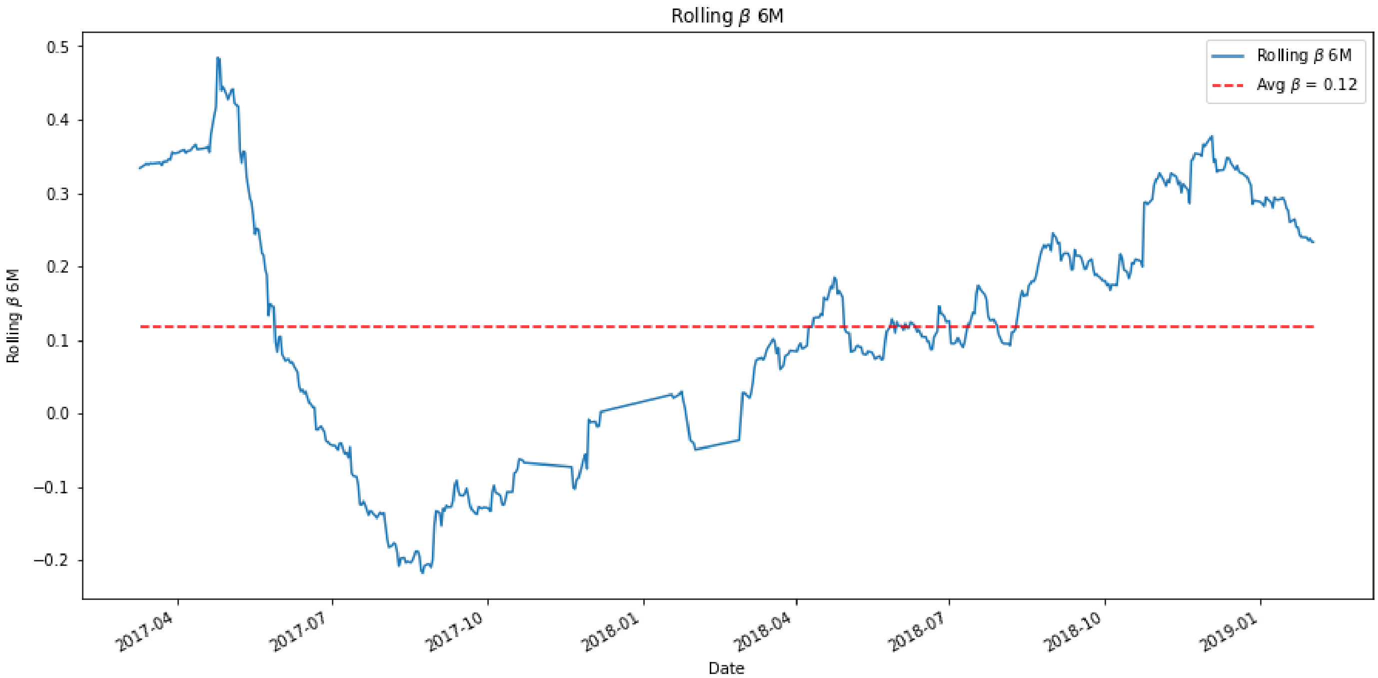

Rolling Annualized α 6M
Rolling α 6M (%)
Avg α = 10.87%
Rolling Annualized α 6M (%)
Date
50
40
30
20
10
0
−10
2010
2011
2012
2013
2014
2015
2016

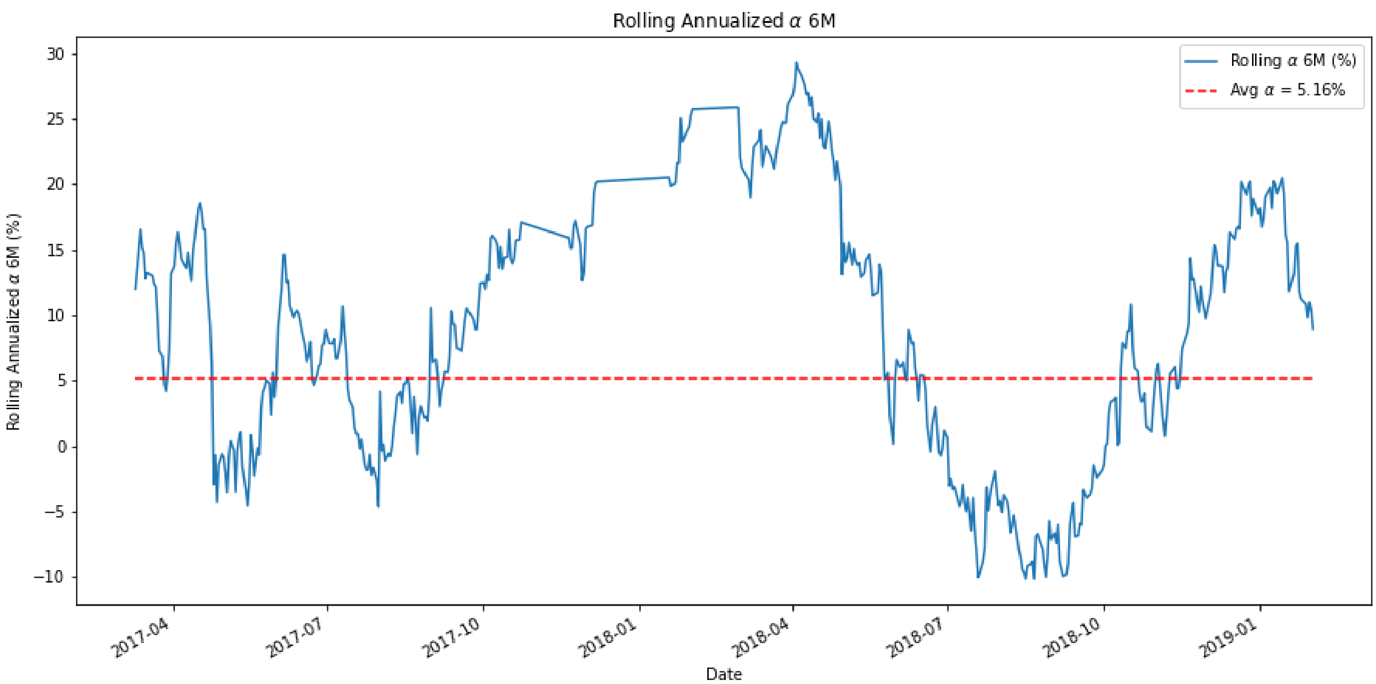

Drawdown (%)
Drawdown (%)
Drawdown (%)
Date
0
−2
−4
−6
−8
−10
2010
2011
2012
2013
2014
2015
2016

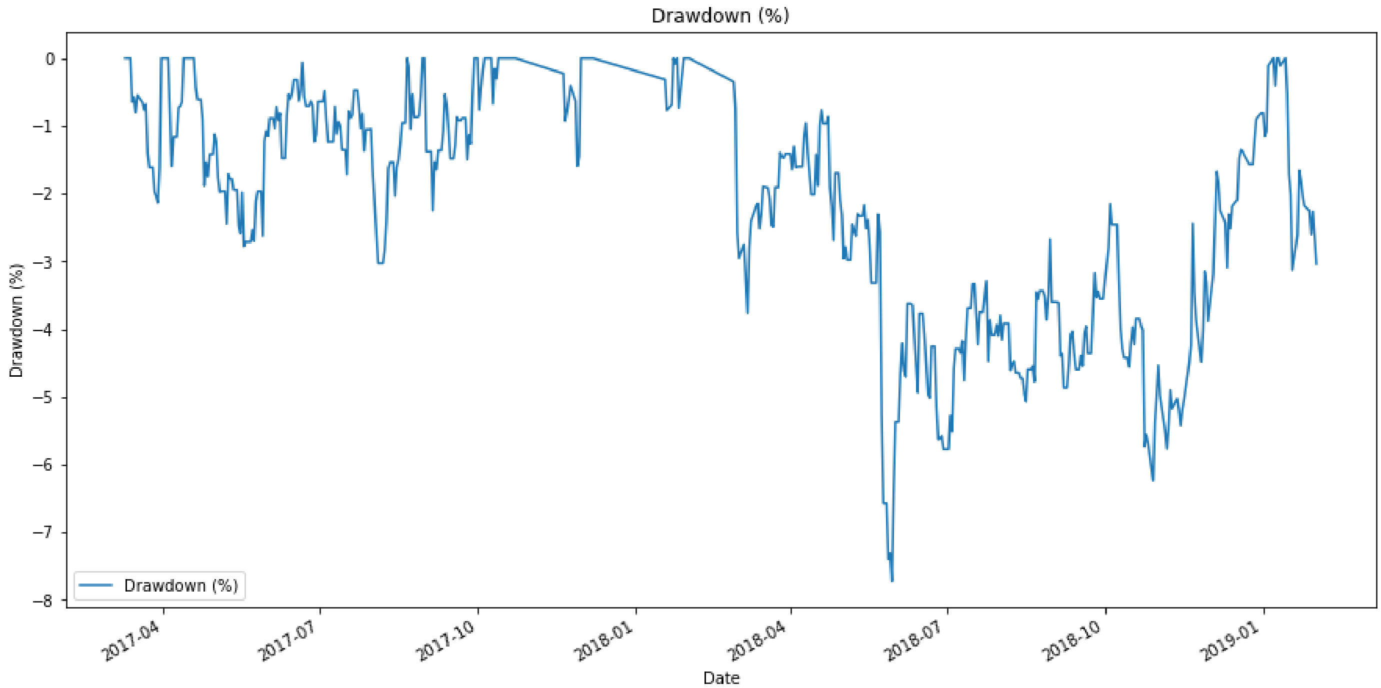


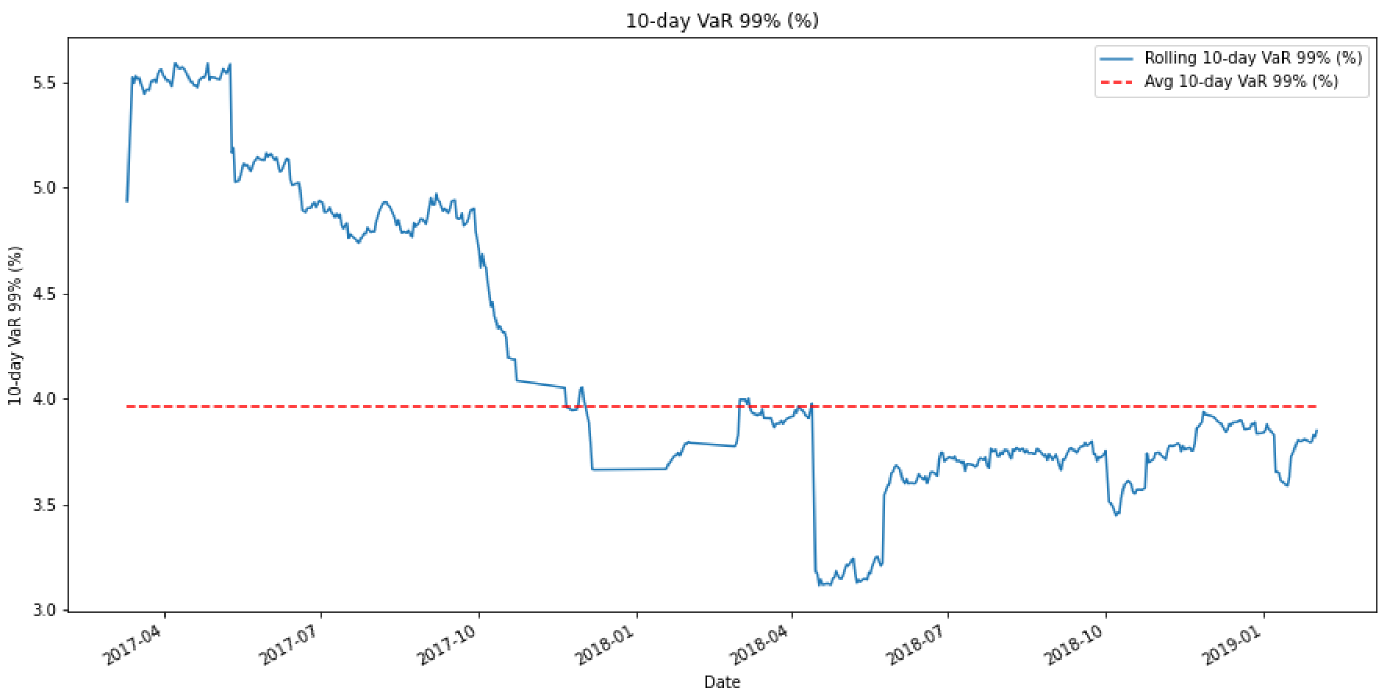
Unsuprinsgly, for the validation period 2017 - 2019 we observed a highly variable Sharpe, Beta and Alpha similar to the training period again due to trades taking too long to be closed making the strategy exposed to market movements




Rolling β 6M
Rolling β 6M
Avg β = 0.12
Rolling β 6M
Date
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
2017-04
2017-07
2017-10
2018-01
2018-04
2018-07
2018-10
2019-01


Rolling Annualized α 6M
Rolling α 6M (%)
Avg α = 5.16%
Rolling Annualized α 6M (%)
Date
30
25
20
15
10
5
0
−5
−10
2017-04
2017-07
2017-10
2018-01
2018-04
2018-07
2018-10
2019-01





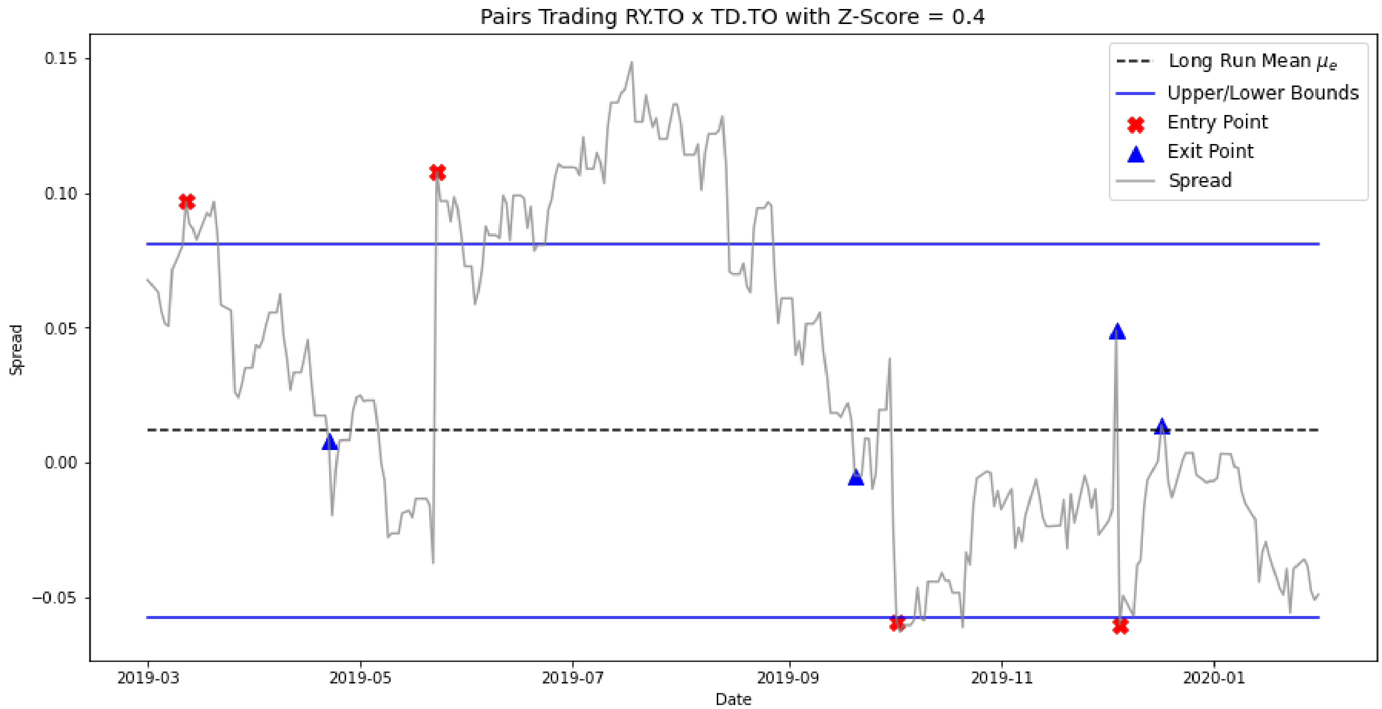
Finally we analyzed the test period between 2019 and 2020 on **Table 10** that sumarizes both P&L and risk-adjusted metrics, and noticed that the Annual return of 4.98% is in line with 2017 - 2019 validation period of 4.50% both at the same volatiliy levels of around 9%. However, the test period alpha is much lower at 1.82% mainly as factor that this experienced a rally driven by tech stocks in the TSX, but its profitability is questionable as the p-value is high. But an interesting point is that the average days a trade is open is now 59 days much closer to the half-life estimated by OU process.

Table 10: RY.TO and TD.TO optimal strategy results : Mar 1st, 2019 - Feb 1st, 2020

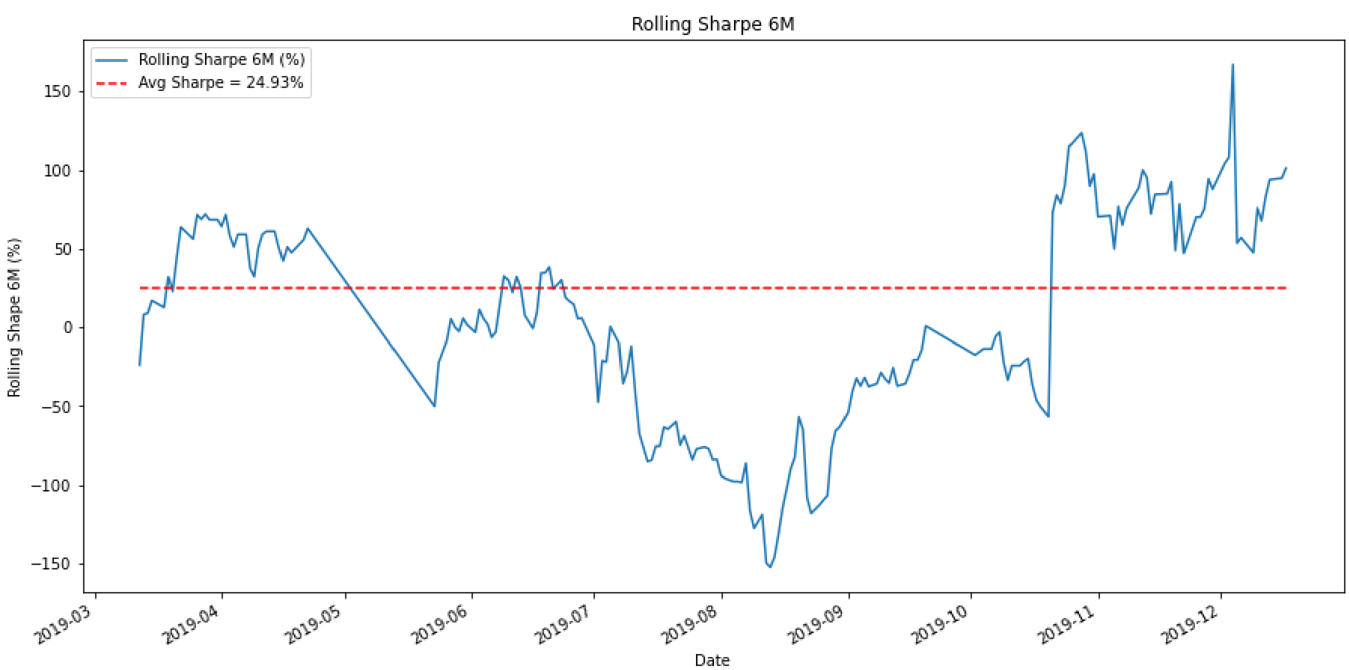
	Strategy ID	Z-Score	# Trades	Annualized Return (%)	Annualized Volatility (%)	Cumulative Return (%)	Avg # Days Open Position	Total Trading Days
0	1.00	0.40	4.00	4.98	9.35	3.83	59.00	236.00

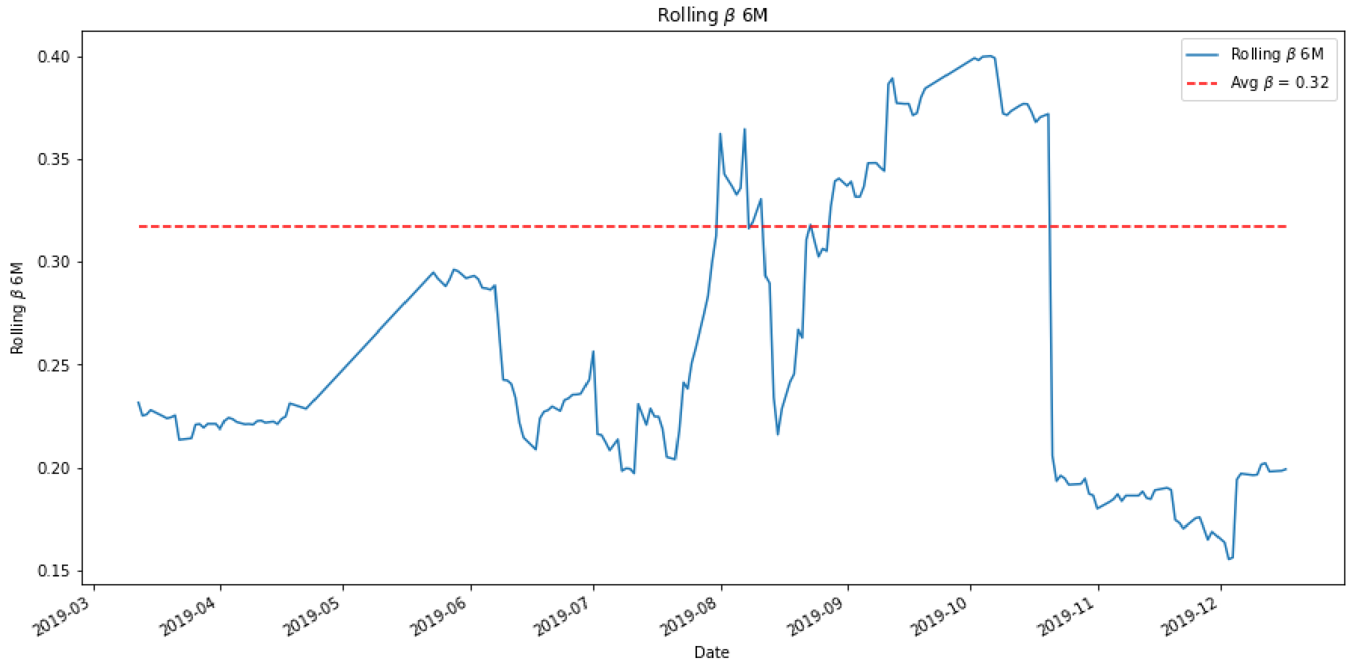
Strategy / Metric	Value (%) Z-Score = 0.4
Annual Return	4.98
Annual Vol	9.35
Cumulative Return	3.83
Alpha (Annual)	1.82
Sharpe	24.93
Beta	31.71
Beta P-Value	0.06
Alpha P-Value	86.09
Max Drawdown	5.85
1-Day VaR 99%	1.35
10-Day VaR 99%	4.12
1-Day ES 99%	1.55
10-Day ES 99%	4.75

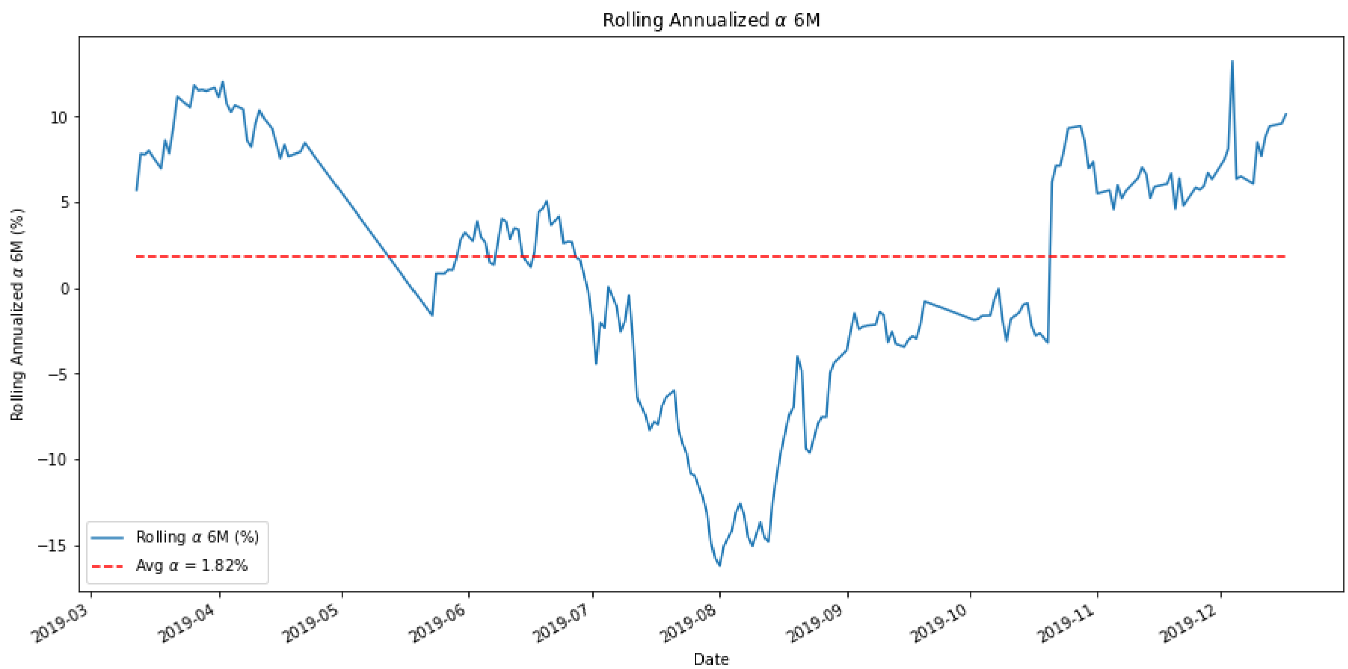
The chart below plots the spread for the testing period between 2019 - 2020 with the Upper/Lower bounds, as well as the four trades entries and exit points. It is clear that the Z=0.4 bounds look fairly wide when compared to the training and validation period, also it seems that there has been a drift in the equilibrium point μ_e indicating that probably a restimation of parameters would be necessary.

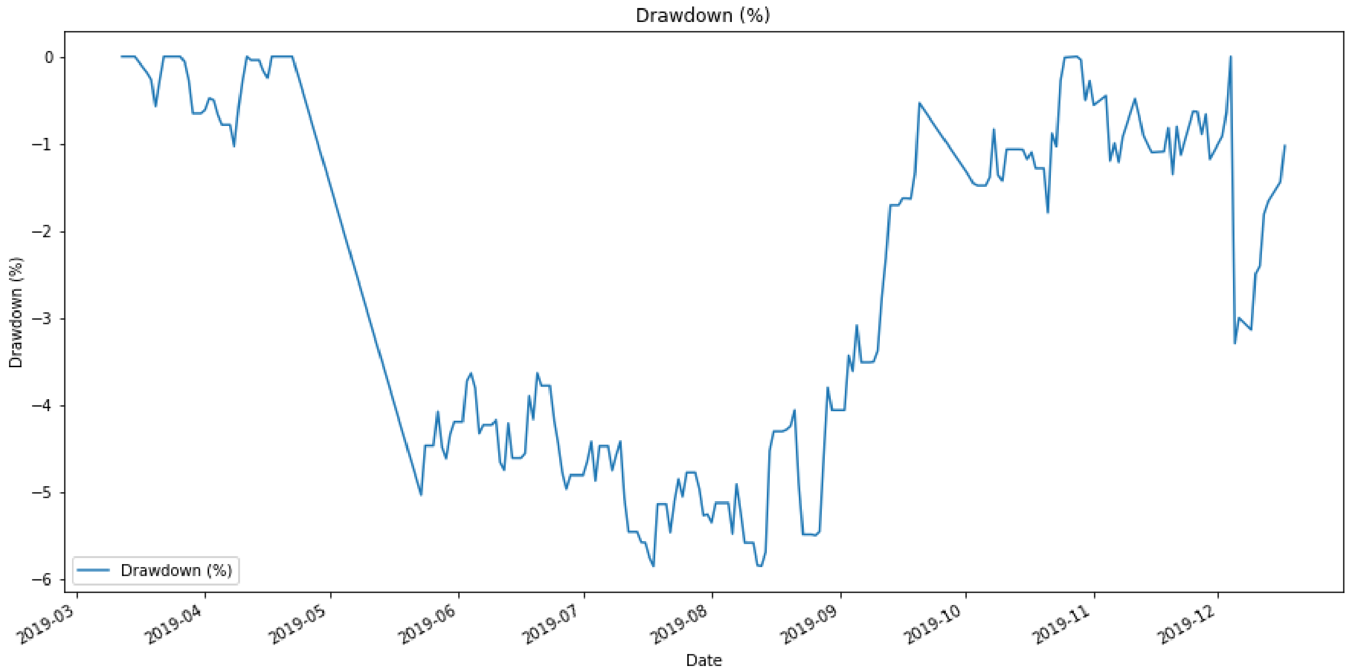


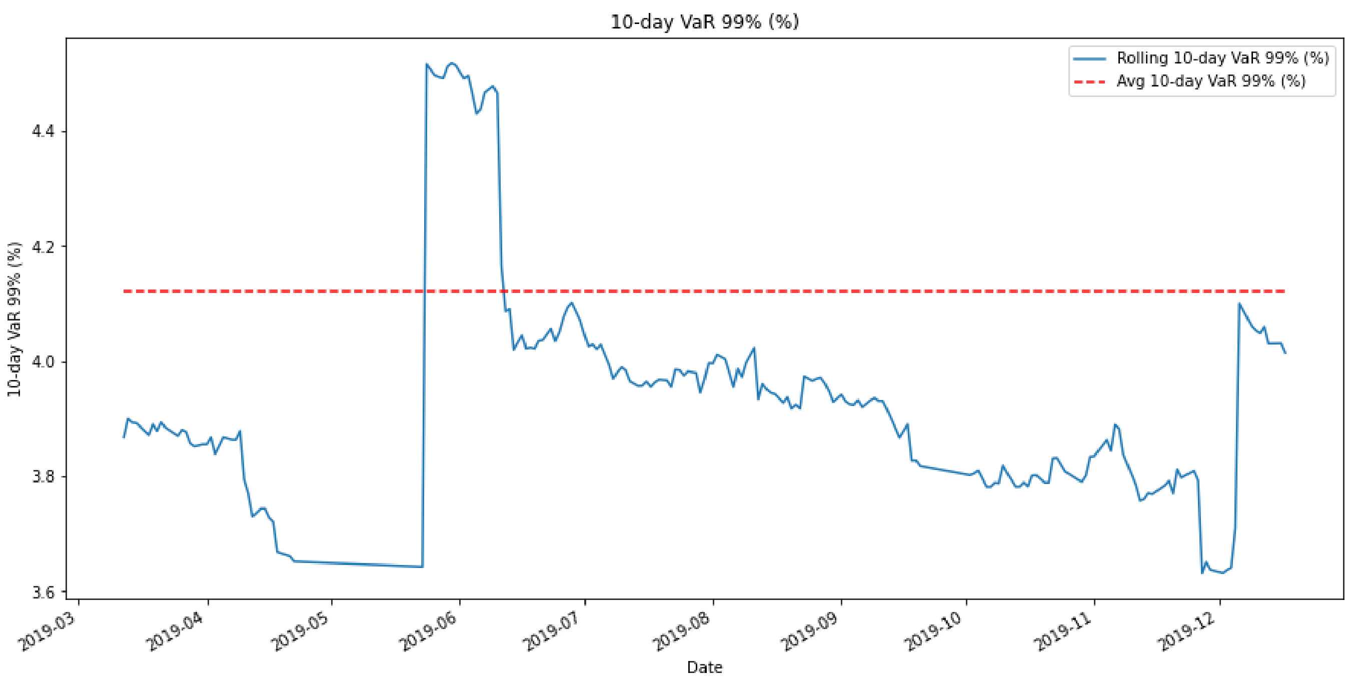
To complete the analysis, the charts below show the rolling statistics and as show previously, all show a high variability around the mean driven especially by the trades that have longer life spans. The focus should be on the rolling alpha chart, which is highly variable confirming that its value of 1.82% for the period might be spurious instead of an actually effective strategy




Rolling β 6M
Rolling β 6M
Avg β = 0.32
Rolling β 6M
0.40
0.35
0.30
0.25
0.20
0.15
2019-03
2019-04
2019-05
2019-06
2019-07
2019-08
2019-09
2019-10
2019-11
2019-12
Date


Rolling Annualized α 6M
Rolling α 6M (%)
Avg α = 1.82%
Rolling Annualized α 6M (%)
10
5
0
−5
−10
−15
2019-03
2019-04
2019-05
2019-06
2019-07
2019-08
2019-09
2019-10
2019-11
2019-12
Date





Pair 2: Japanese Yen (JPYUSD) x Gold Futures (GC=F)

Engle Granger Fitting

Table 11 shows the two linear regression ran for this par, and as expected all the coefficients are highly significant but are spurious, which is not enough to conclude if this is an appropriate pair for long/short strategy.

Table 11: JPYUSD and Gold OLS Regression Results

	Variable				GC=F	Variable				JPYUSD=X
Statistic		**Beta**	**P-Value**	**SE**	**T-Stat**		**Beta**	**P-Value**	**SE**	**T-Stat**
0	Constant	0.279	0.000	0.016	17.247	Constant	0.178	0.000	0.014	13.041
1	JPYUSD=X	0.966	0.000	0.016	58.865	GC=F	0.652	0.000	0.011	58.865

Table 12 summarizes the main ADF statistics for the residual/spreads of both equations, and the results are very interesting as Equation 1 residuals is significant at the 5% i.e. the spread is stationary, on the other hand, when we reverse the regression the residuals are not significant at the 5% level. This is potentially a red flag that maybe the spread is not really stationary and the cointegration between these two series is weak. It would be fair to expect similar T-statistic, and not so wildly different.
If it were the case to implement this strategy I would stop the analysis here and not proceed with this pair, however for the sake of the project there is enough evidence of stationariety in Equation 1

Table 12: JPYUSD and Gold ADF Test Results

	Equation 1:	**GC=F = JPYUSD=X + Constant**		**Equation 2:**	**JPYUSD=X = GC=F + Constant**	
Null Hypothesis	Series has unit root			Series has unit root		
T-Statistic	-3.394434			-2.826998		
P-Value (MacKinnon)	0.011150			0.054520		
Optimal Lag	24			24		
Confidence Level	1%	5%	10%	1%	5%	10%
Mackinnon Critical Value	-3.433569	-2.862962	-2.567527	-3.433569	-2.862962	-2.567527
Reject/Not Reject H0	Not Reject	Reject	Reject	Not Reject	Not Reject	Reject
Stationary/Non Stationary	Non Stationary	Stationary	Stationary	Non Stationary	Non Stationary	Stationary

Table 13 is the error correction model for Equation 1, again the $1 - \alpha$ is highly significant, hence we can move to estimating the OU process parameters.

Table 13: Equation 1: Gold = JPYUSD + Const Error Correction Model Results

				dY
Statistic	**Beta**	**P-Value**	**SE**	**T-Stat**
Variable				
dX	0.158	0.002	0.050	3.137
Residual(-1)	-0.010	0.000	0.003	-3.918

Table 14 exhibits the OU process parameters and the spread is around 1.23% and half-life is 76 days, hence we can this a strategy that will have positions opens for longer periods of time and more exposed to market movements, being overall more risky.

Table 14: JPYUSD and Gold OU Process Fitted Parameters

	Value
OU Process Parameters	
μ_e	0.012282
Half-Life (days)	76.024781
σ	0.290711
θ	2.297581

The chart below shows in blue the spread between gold and Japanese Yen, and the read line if the equilbrium point μ_e, it is interesting to note that the series has a clear upward trend between 2010 and 2012, which then stabilizes afterwards. However the period between 2012 - 2014 there is a lot higher variability than 2015-2019 potentially explaining why the T-statistic results are divergent when we fit the two equations.

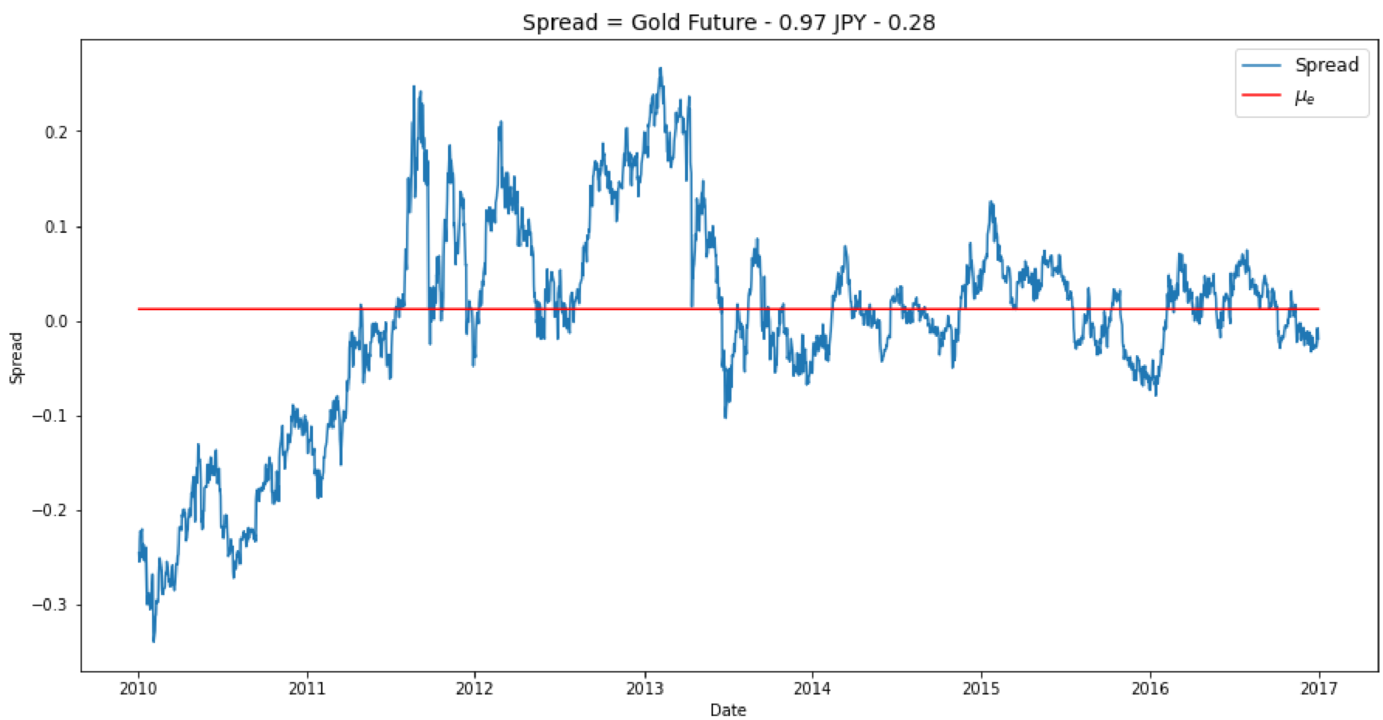


Table 15 exhibits the P&L results for several Z-scores and the most interesting point is that for Z-score = 0.2 the average lifespance of a trade is very similar to the computed OU Half-life, this trategy producres the highest cumulative return with 20 trades. And as expected as we increase the Z-score the number of trade

reduces, average lifespan of trades increase and there is non-linear relationship between returns and volatility.

Table 15: JPYUSD and Gold P&L Results: Jan 1st, 2010 - Dec 31st, 2016

	Strategy ID	Z-Score	# Trades	Annualized Return (%)	Annualized Volatility (%)	Cumulative Return (%)	Avg # Days Open Position	Total Trading Days
0	1.00	0.20	20.00	17.69	19.47	119.35	76.65	1533.00
1	2.00	0.30	9.00	20.45	20.34	84.68	117.44	1057.00
2	3.00	0.40	6.00	18.35	20.25	62.81	154.00	924.00
3	4.00	0.50	5.00	20.38	20.25	63.28	169.40	847.00
4	5.00	0.60	4.00	20.19	20.13	57.09	197.25	789.00
5	6.00	0.70	3.00	25.00	21.00	52.42	200.67	602.00

Table 16 exhibits risk-adjusted metrics (in percentage terms) for several Z-scores . Again, it is interesting to note that Z-score = 0.2 provides the results most aligned with the intuition and the goal of this long/short strategy which is to be market-neutal, the beta is -0.35% even though is not statistically significant it provides a interesting intuition, on the other hand Z score= 0.3 is also very low, only -10% and highly significant statistically. Based on this table, the best strategy seems to Z-score = 2 as it is the only one that provide statistically significant alpha over the Nikkei 225.

Table 16: JPYUSD and Gold Risk-Return Metrics: Jan 1st, 2010 - Dec 31st, 2016

						Value (%)
Strategy	**Z-Score = 0.2**	**Z-Score = 0.3**	**Z-Score = 0.4**	**Z-Score = 0.5**	**Z-Score = 0.6**	**Z-Score = 0.7**
Metric						
Annual Return	17.69	20.45	18.35	20.38	20.19	25.00
Annual Vol	19.47	20.34	20.25	20.25	20.13	21.00
Cumulative Return	119.35	84.68	62.81	63.28	57.09	52.42
Alpha (Annual)	18.18	20.74	18.33	19.94	20.54	24.42
Sharpe	74.68	83.01	74.29	82.41	81.91	97.37
Beta	-0.36	-10.75	-10.23	-6.57	2.55	17.99
Beta P-Value	93.23	4.11	7.60	27.40	69.32	1.63
Alpha P-Value	4.07	6.39	12.38	11.02	11.07	10.89
Max Drawdown	18.21	15.59	14.50	14.34	14.34	12.79
1-Day VaR 99%	2.78	2.90	2.89	2.89	2.87	2.98
10-Day VaR 99%	8.30	8.61	8.64	8.57	8.52	8.76
1-Day ES 99%	3.20	3.33	3.33	3.32	3.30	3.43
10-Day ES 99%	9.62	9.98	10.00	9.93	9.88	10.18

Table 17 and Table 18 show, respectively, P&L and risk-return metrics for the validation period between 2017 - 2019 annd the first point is that only the first two strategies (Z-score equal 0.2 and 0.3) had trades executed. This is a major red flag, as there has been a drift in the equilibrium point μ_e which basically nullified the strategy. Note that Z-score = 0.2 had only two trades and a measly return of 1.45%, Z-score = 0.3 had a better annualized return but only one trade.
In order to maintain the consistency in the selection criteria I pick Z-score = 0.3 because it had the largest out-of-sample alpha of 7.03%

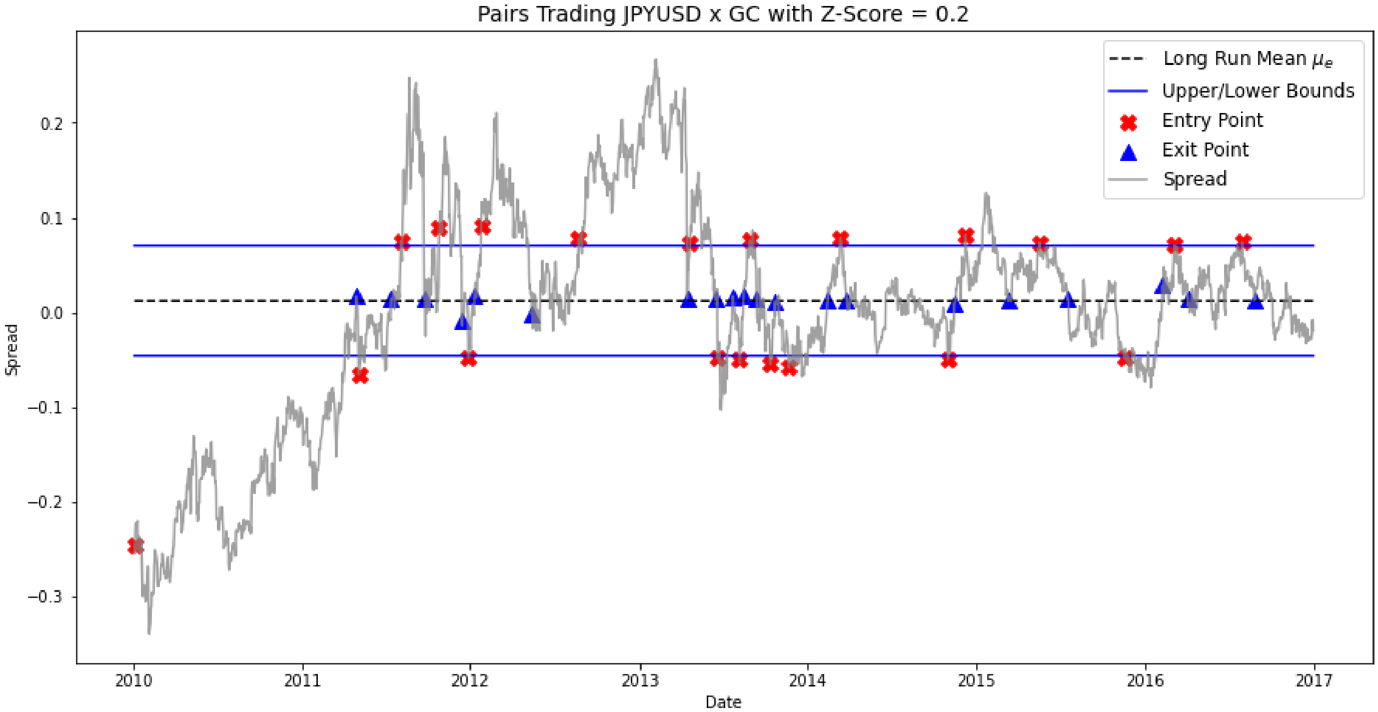
Table 17: JPYUSD and Gold P&L Results: Feb 1st, 2017 - Feb 1st, 2019

	Strategy ID	Z-Score	# Trades	Annualized Return (%)	Annualized Volatility (%)	Cumulative Return (%)	Avg # Days Open Position	Total Trading Days
0	1.00	0.20	2.00	1.45	10.13	1.53	166.00	332.00
1	2.00	0.30	1.00	6.05	10.25	6.03	318.00	318.00

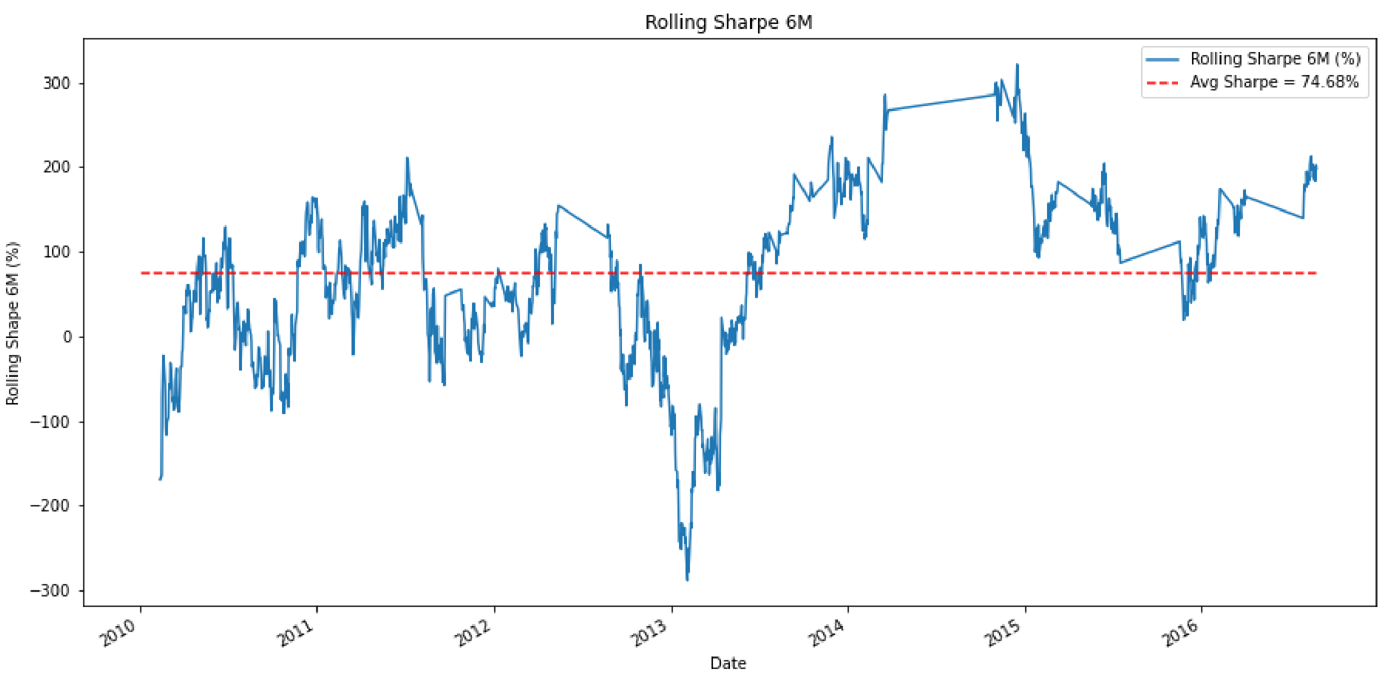
Table 18: JPYUSD and Gold Risk-Return Metrics: Feb 1st, 2017 - Feb 1st, 2019

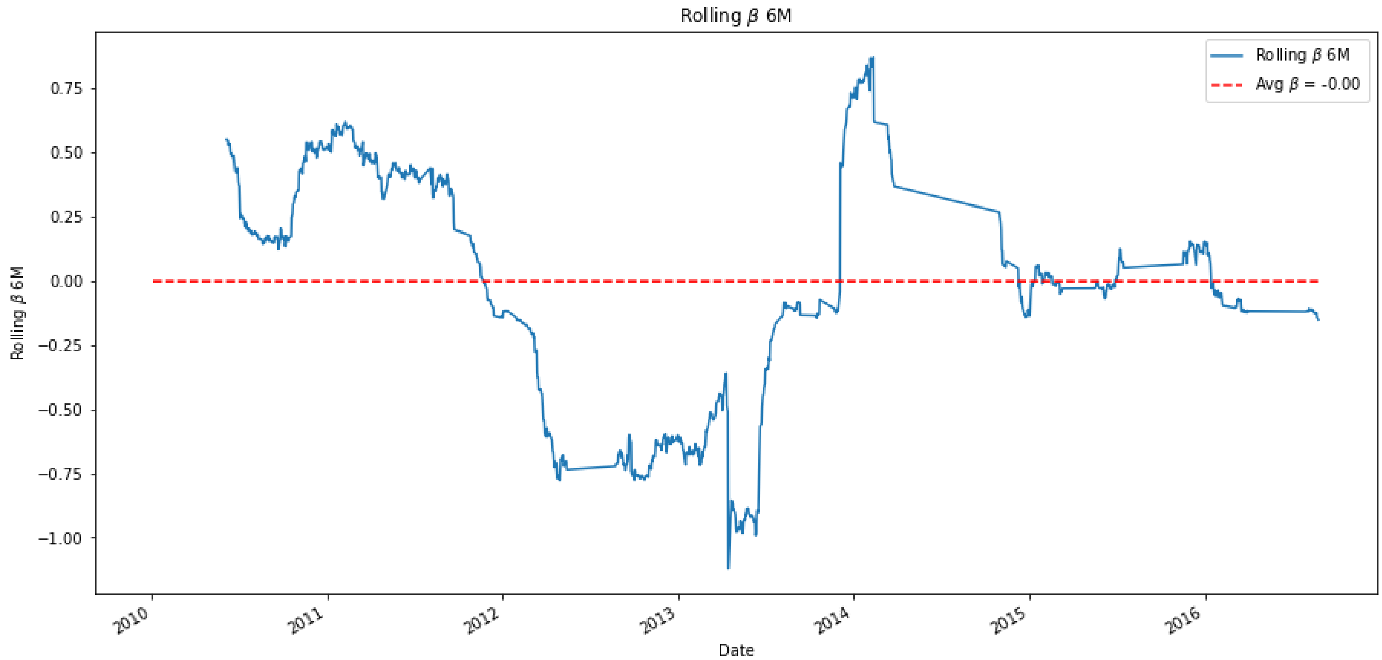
	Value (%)	
Strategy	**Z-Score = 0.2**	**Z-Score = 0.3**
Metric		
Annual Return	1.45	6.05
Annual Vol	10.13	10.25
Cumulative Return	1.53	6.03
Alpha (Annual)	2.60	7.03
Sharpe	-8.94	34.55
Beta	-7.95	-7.44
Beta P-Value	32.07	36.39
Alpha P-Value	79.31	49.57
Max Drawdown	6.11	6.11
1-Day VaR 99%	1.48	1.48
10-Day VaR 99%	4.62	4.50
1-Day ES 99%	1.69	1.70
10-Day ES 99%	5.30	5.19

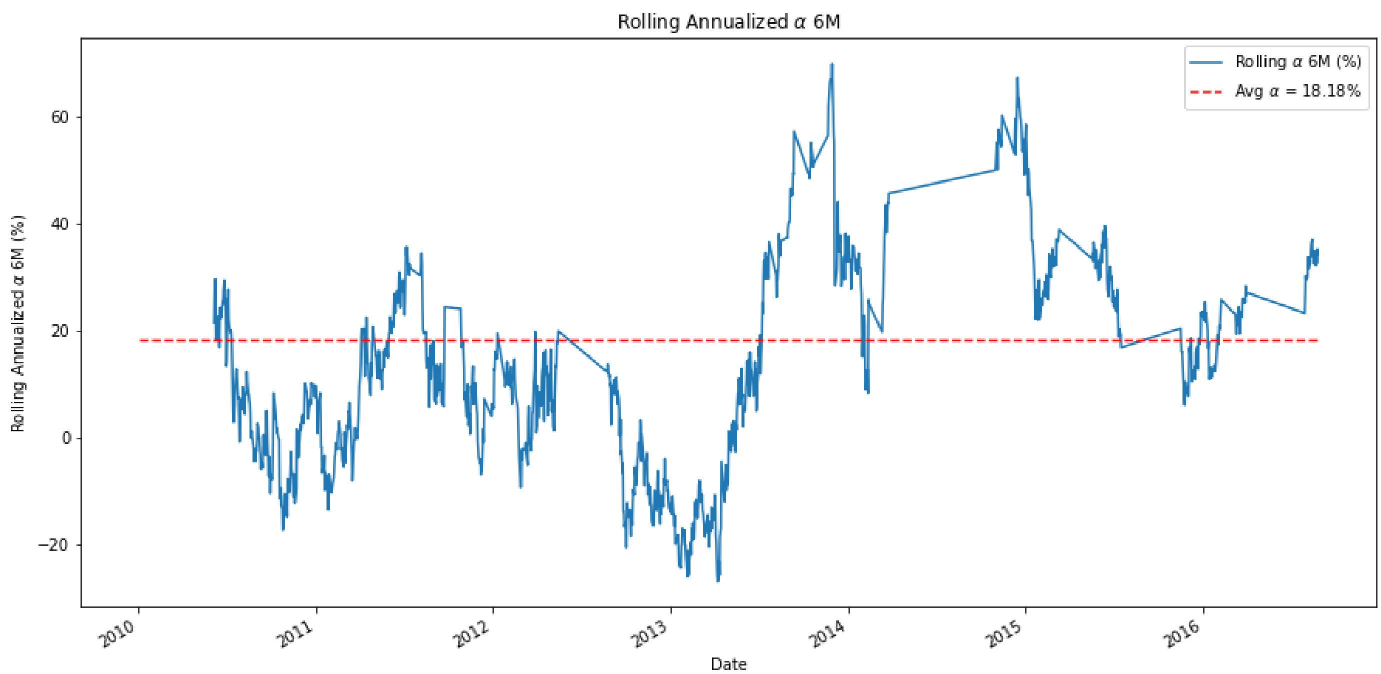
The chart below plots the entry and exit points for the twenty trades during the training period of the strategy (2010-2017) and we can see that there a several trades occuring in a short period of time between 2013 and 2015, and then more sparsely from 2015 onwards - this pattern of trades is what is expected for this time of strategy.

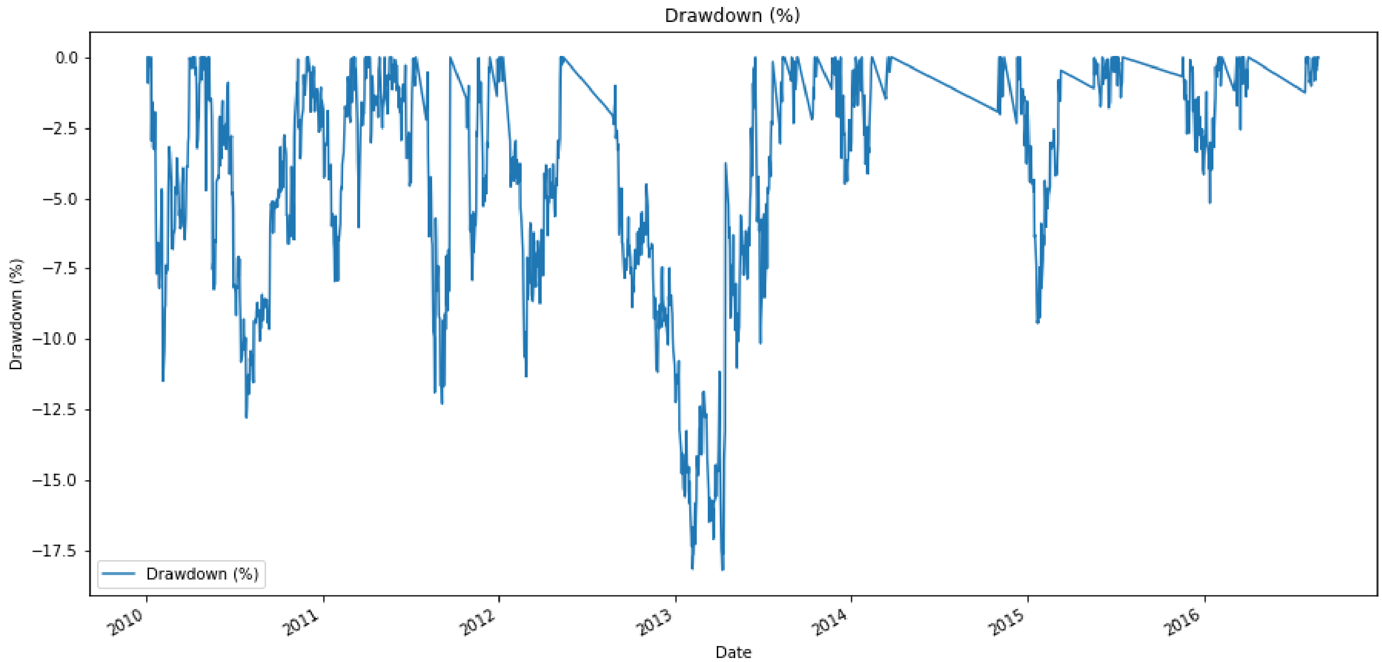


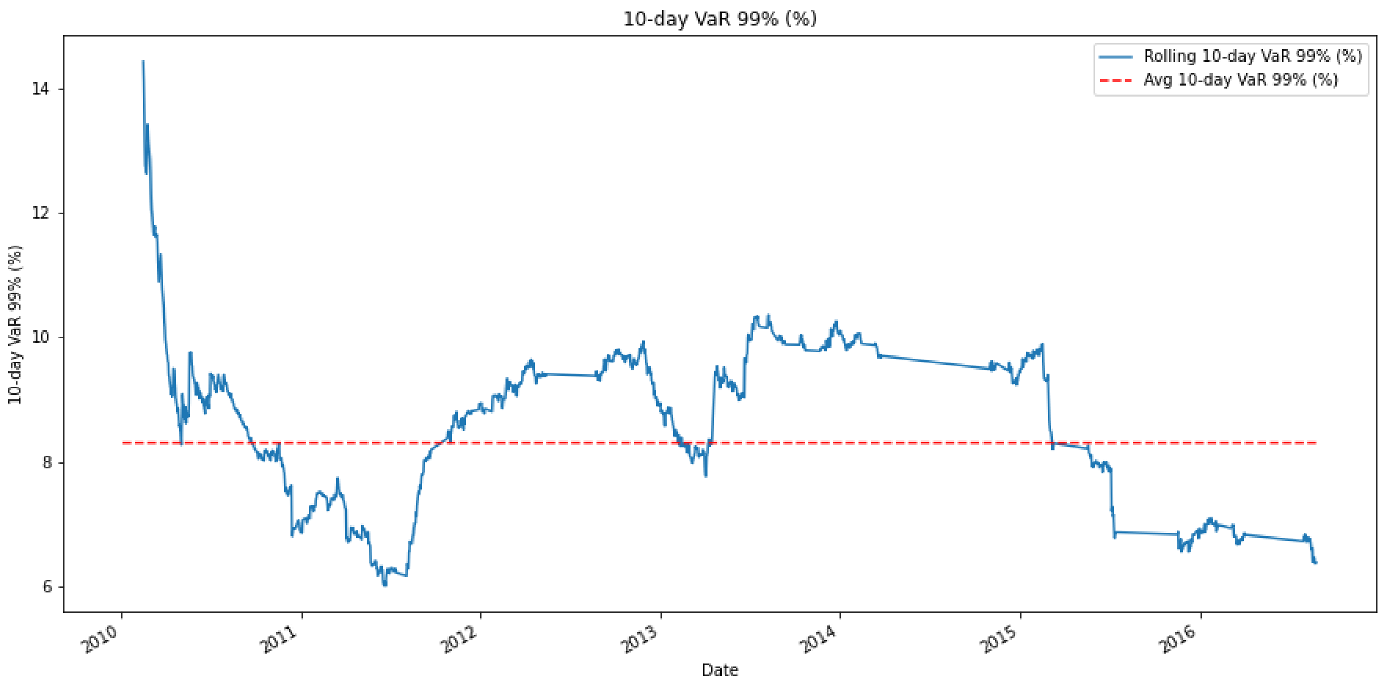
The following four charts plot rolling six month risk adjusted metrics for the optimal strategy with Z-score = 0.2. The main takeaway is that as observed with pair 1 both Sharpe and Beta oscillate around the mean due to especially long lifespan trades that keep the position exposed to market risk.



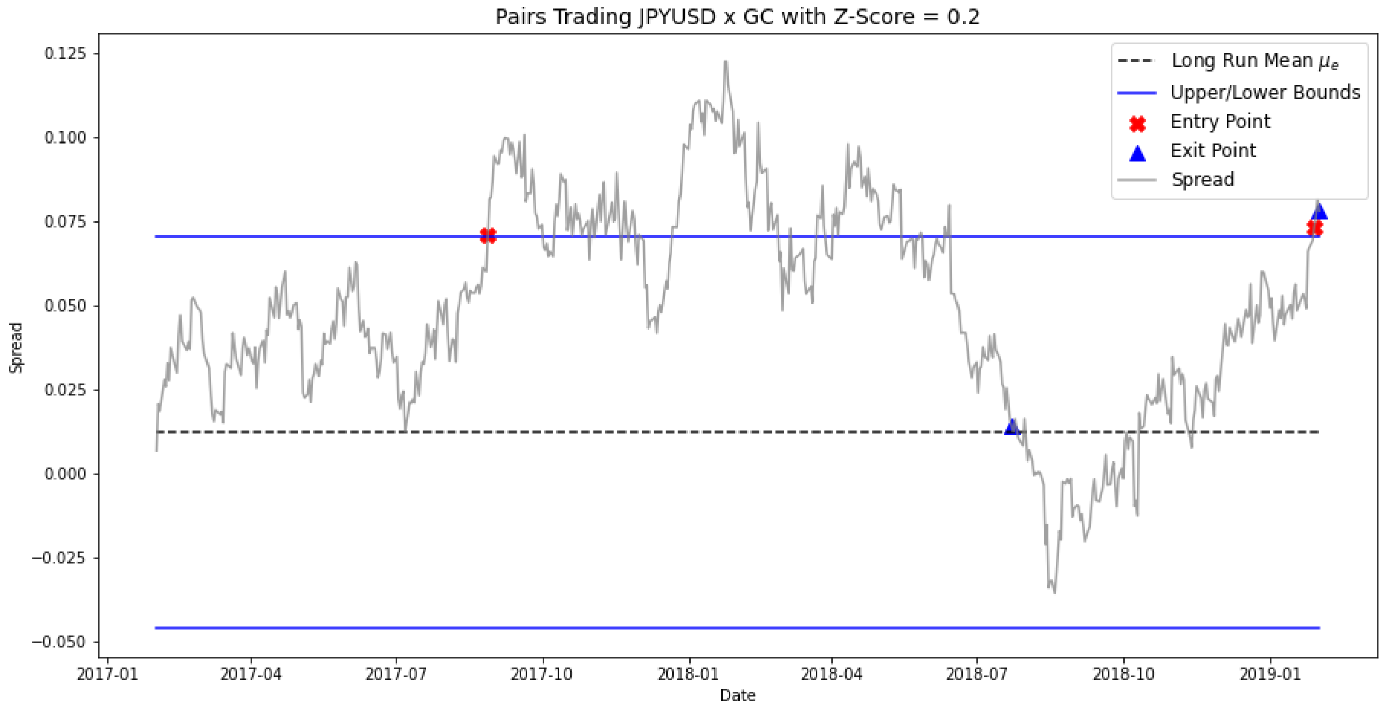

Rolling β 6M
Rolling β 6M
Avg β = -0.00
Rolling β 6M
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
2010
2011
2012
2013
2014
2015
2016
Date


Rolling Annualized α 6M
Rolling α 6M (%)
Avg α = 18.18%
Rolling Annualized α 6M (%)
60
40
20
0
−20
2010
2011
2012
2013
2014
2015
2016
Date

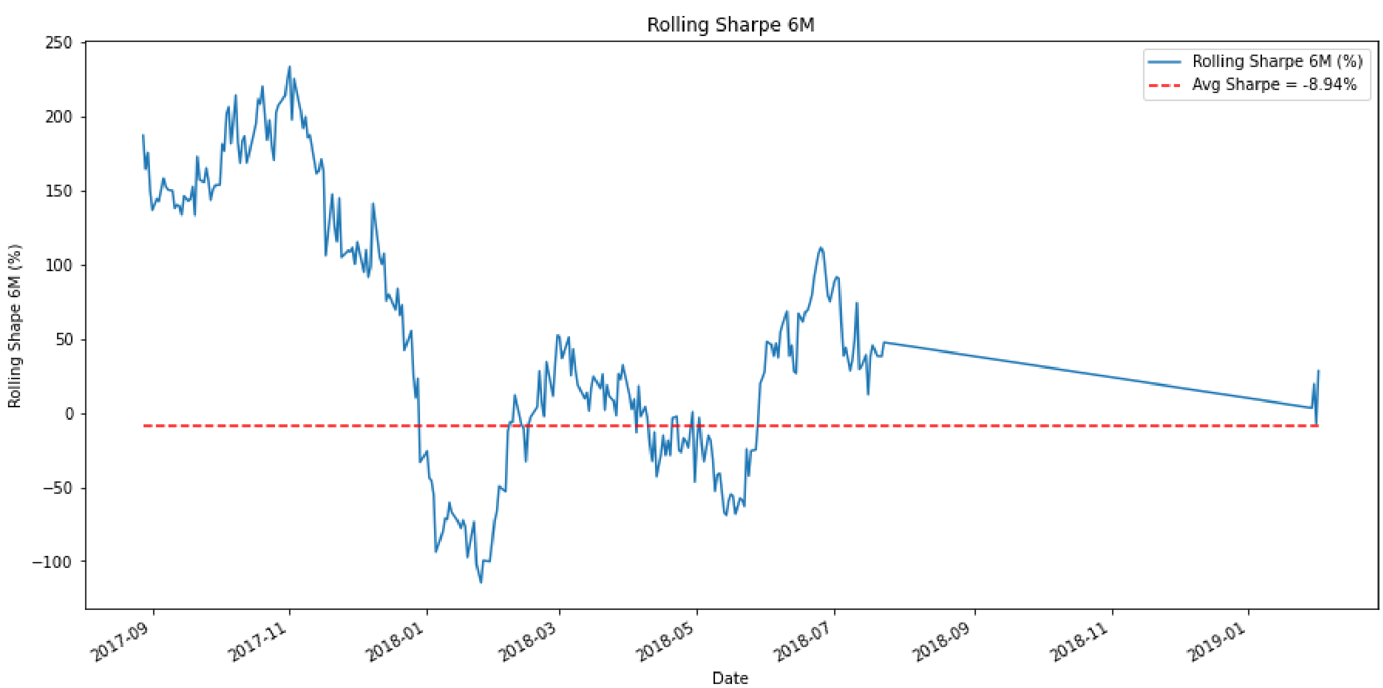


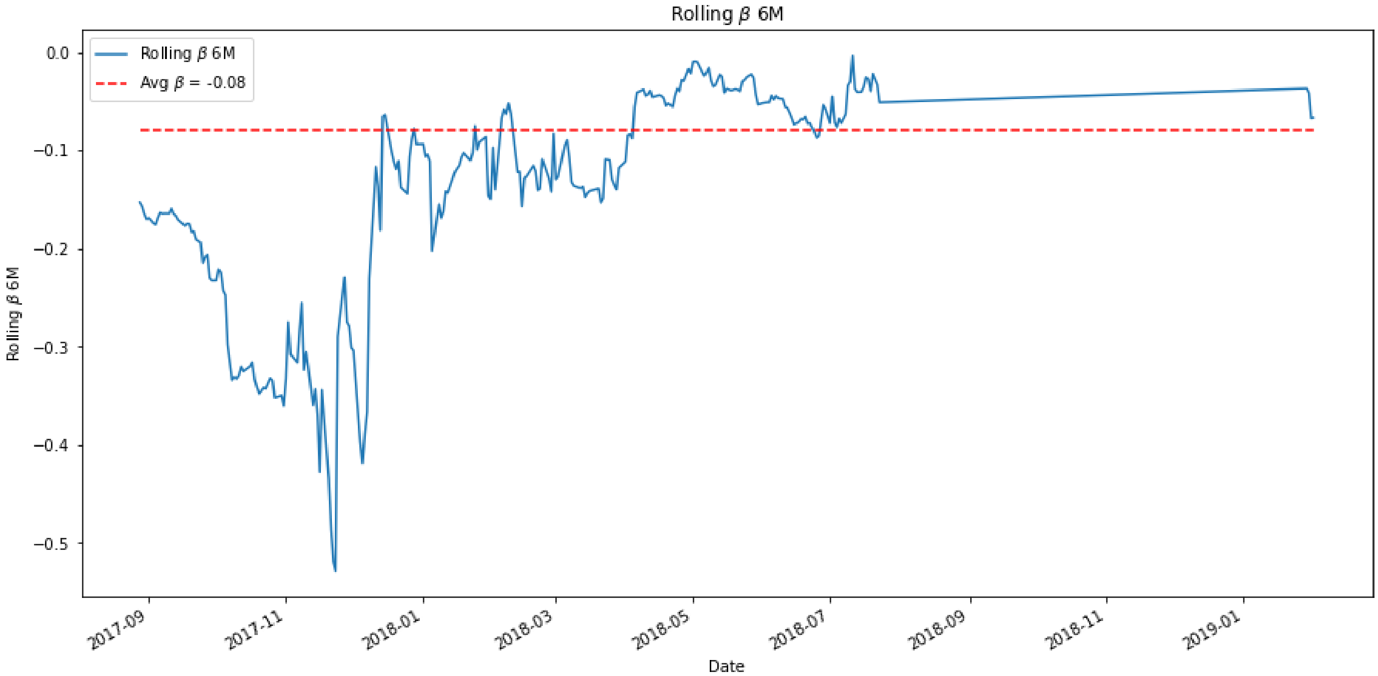


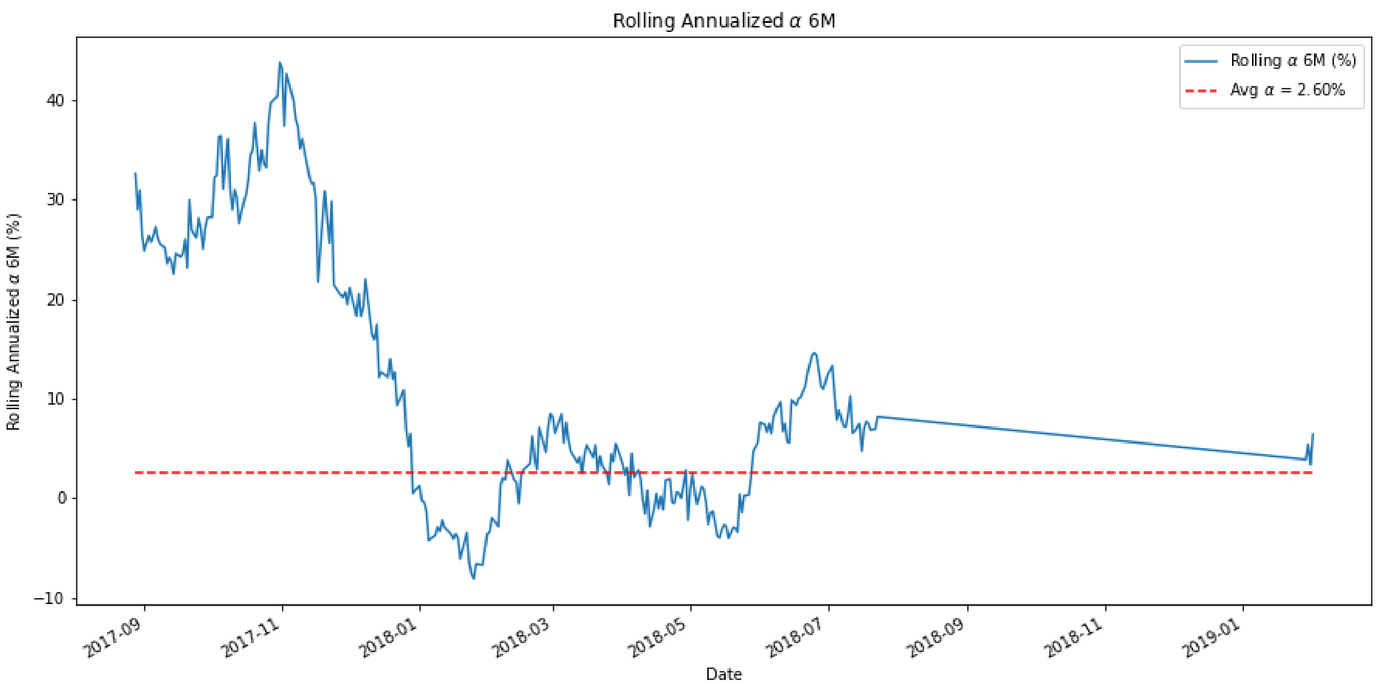
The figure below is trading pattern for the validation period with the optimal Z-score = 0.2; Note that only two trades were executed. The second one is very short-lived, opening and being closed because the trading period is over. But it seems that there was a drift in equilibrium point mu_e, because between the interval between 2017-07 and 2018-07 is around 252 days, about times the expected half-life of 76 days estimated by the OU process

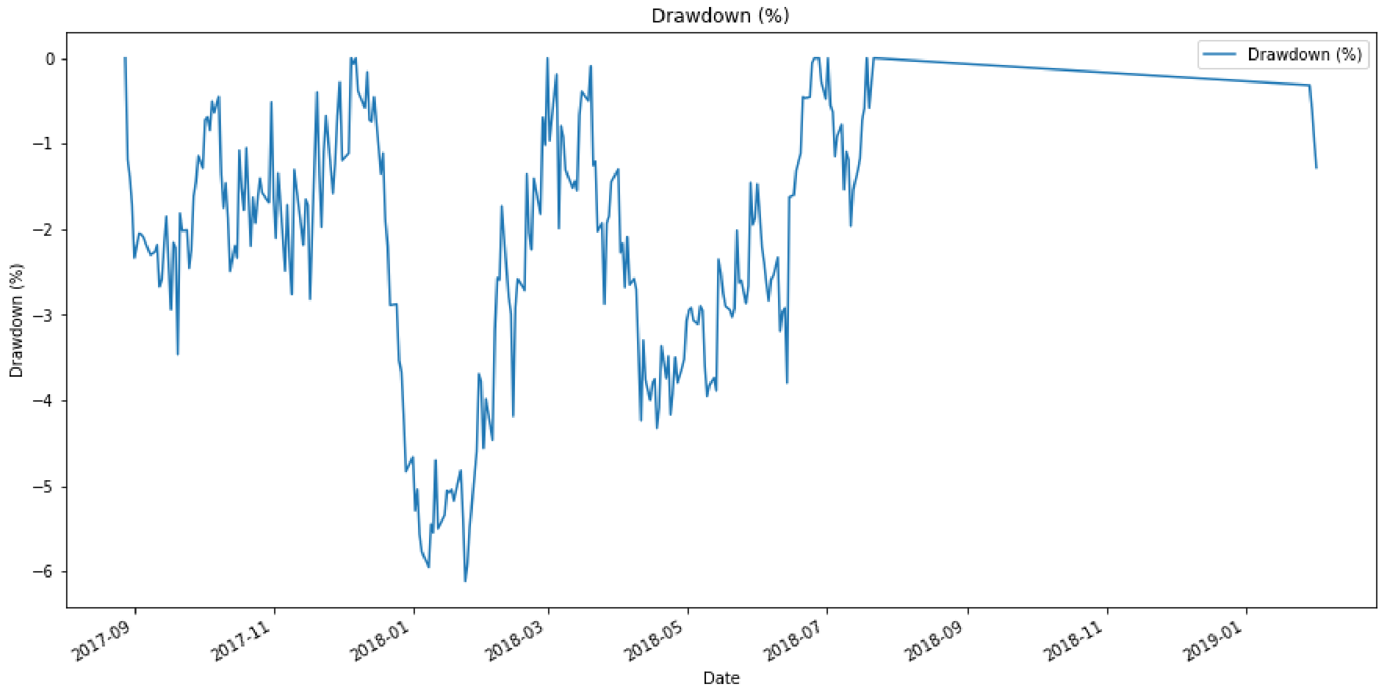


The next four charts plot the rolling risk metrics for the validation period, for this analysis due to the small number of trades it is very sparse and the charts are rather meaningless because of the rolling nature they carry a lot of information for past training results, they show clearly that the strategy fails out-of-sample




Rolling β 6M
Rolling β 6M
Avg β = -0.08
Rolling β 6M
Date
0.0
−0.1
−0.2
−0.3
−0.4
−0.5
2017-09
2017-11
2018-01
2018-03
2018-05
2018-07
2018-09
2018-11
2019-01


Rolling Annualized α 6M
Rolling α 6M (%)
Avg α = 2.60%
Rolling Annualized α 6M (%)
Date
40
30
20
10
0
−10
2017-09
2017-11
2018-01
2018-03
2018-05
2018-07
2018-09
2018-11
2019-01



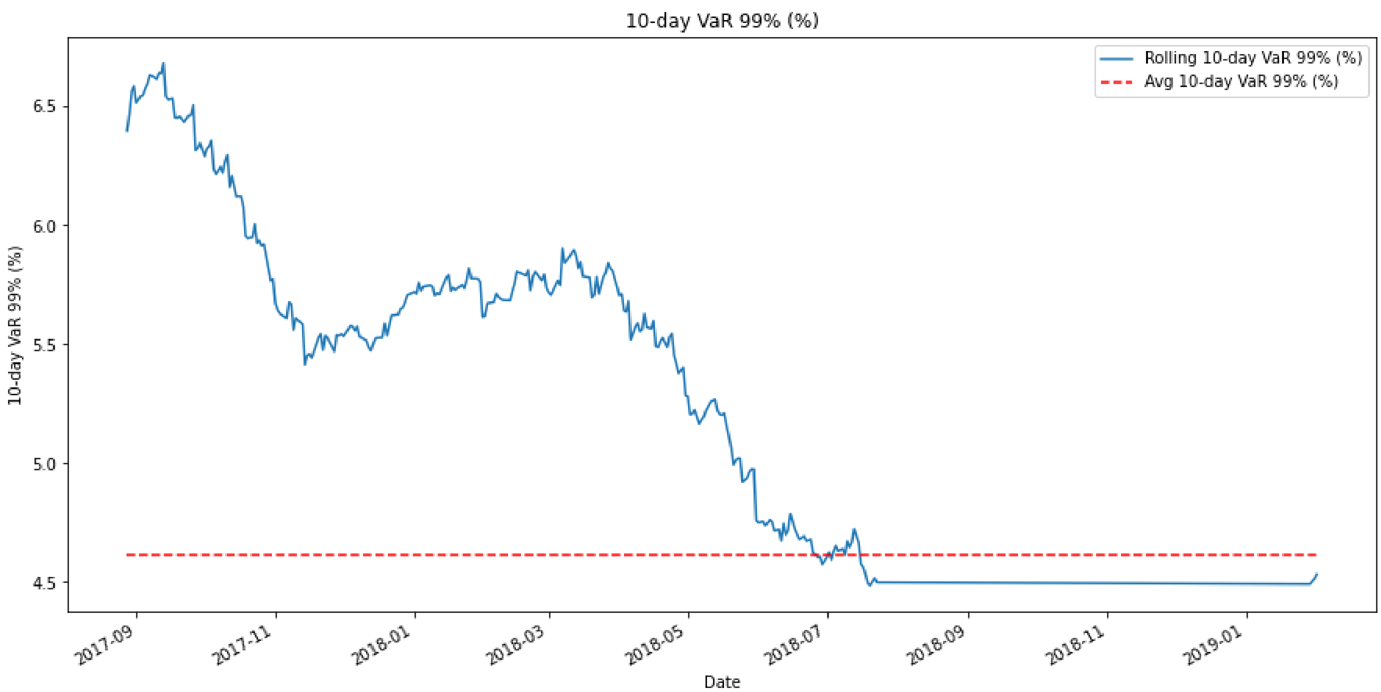


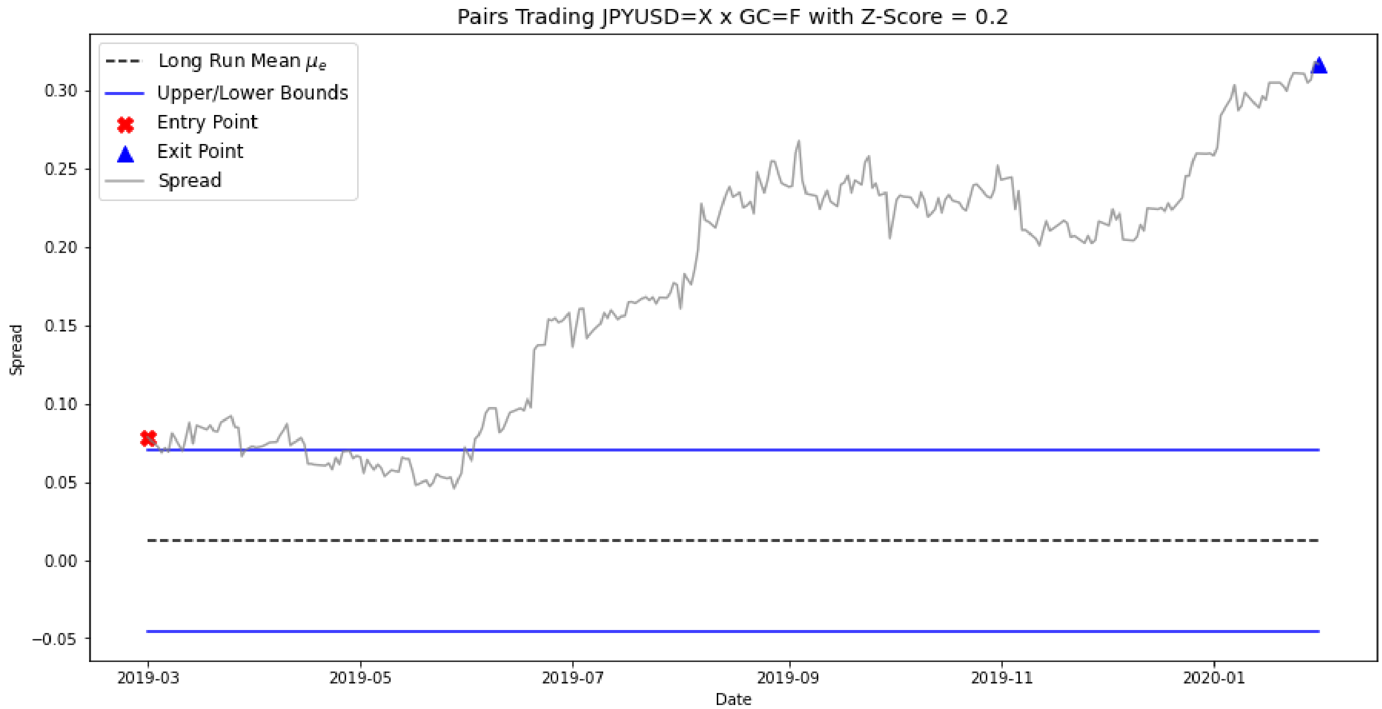
Table 19 summarizes the P&L and risk-adjusted metrics for the training sample and it is very clear from the annualized return of -16.38% that the strategy failed miserably, confirming the early hypothesis that the diverging T-statistic for ADF test might be an indicator of weak cointegration.

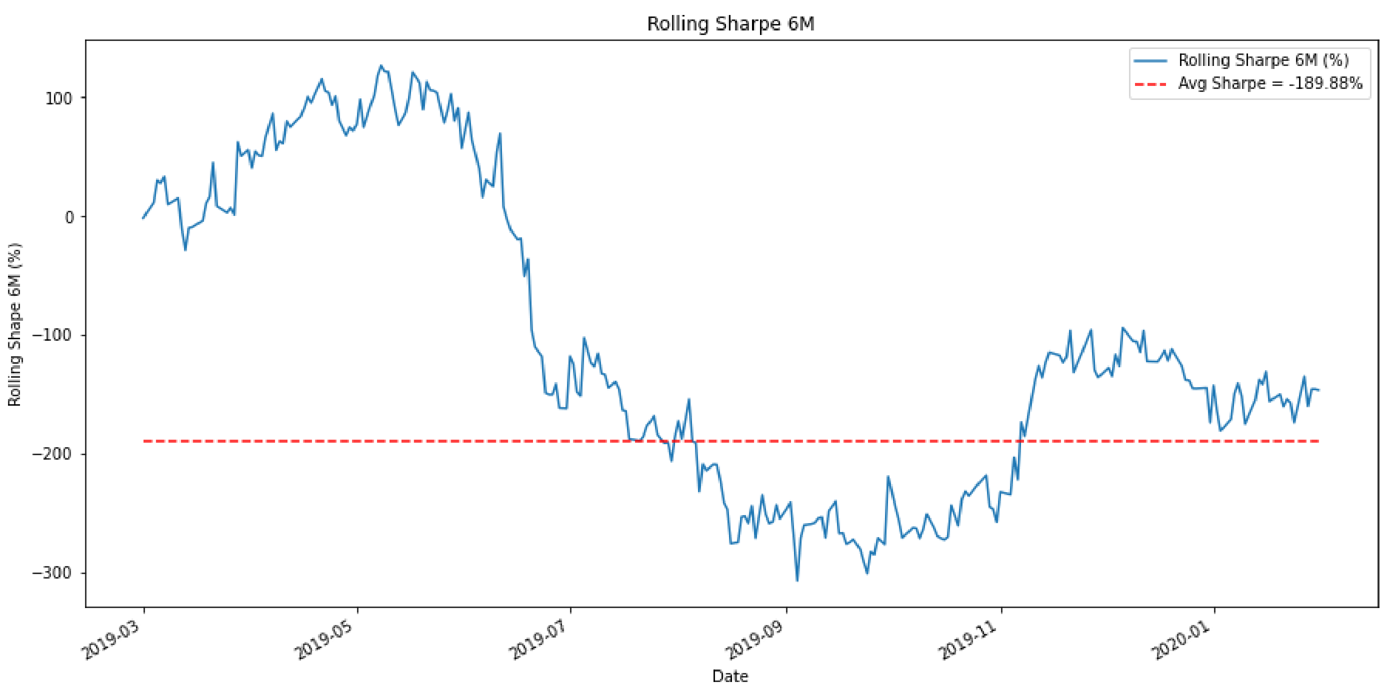
Table 19: JPYUSD and Gold optimal strategy results : Mar 1st, 2019 - Feb 1st, 2020

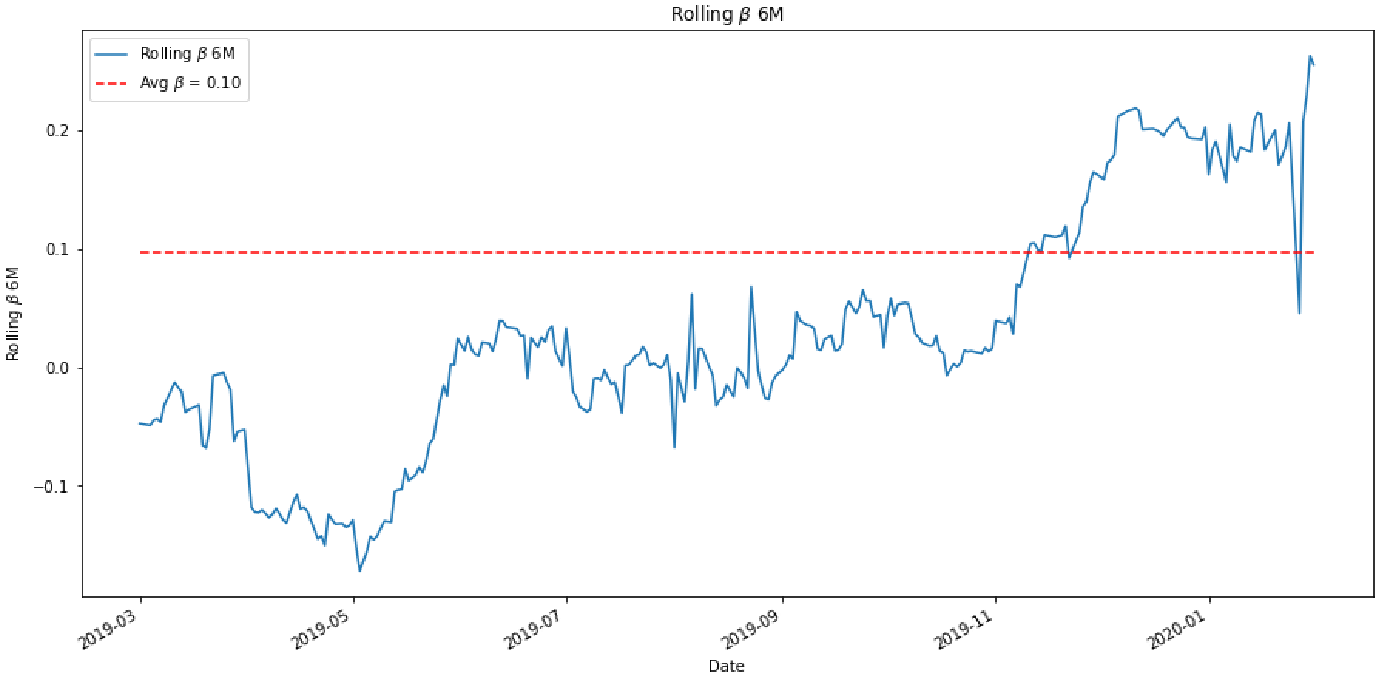
	Strategy ID	Z-Score	# Trades	Annualized Return (%)	Annualized Volatility (%)	Cumulative Return (%)	Avg # Days Open Position	Total Trading Days
0	1.00	0.20	1.00	-16.38	10.67	-17.44	336.00	336.00

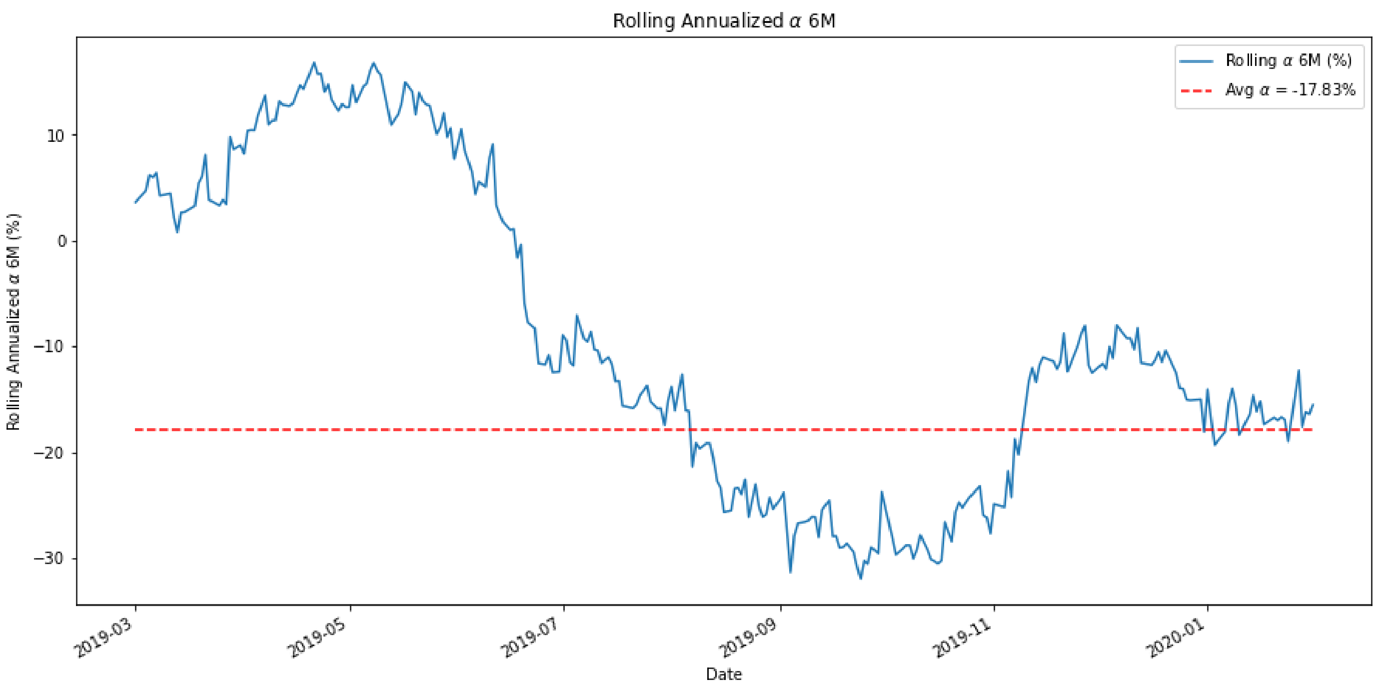
	Value (%)
Strategy	**Z-Score = 0.2**
Metric	
Annual Return	-16.38
Annual Vol	10.67
Cumulative Return	-17.44
Alpha (Annual)	-17.83
Sharpe	-189.88
Beta	9.71
Beta P-Value	30.09
Alpha P-Value	8.53
Max Drawdown	19.99
1-Day VaR 99%	1.63
10-Day VaR 99%	5.63
1-Day ES 99%	1.86
10-Day ES 99%	6.35

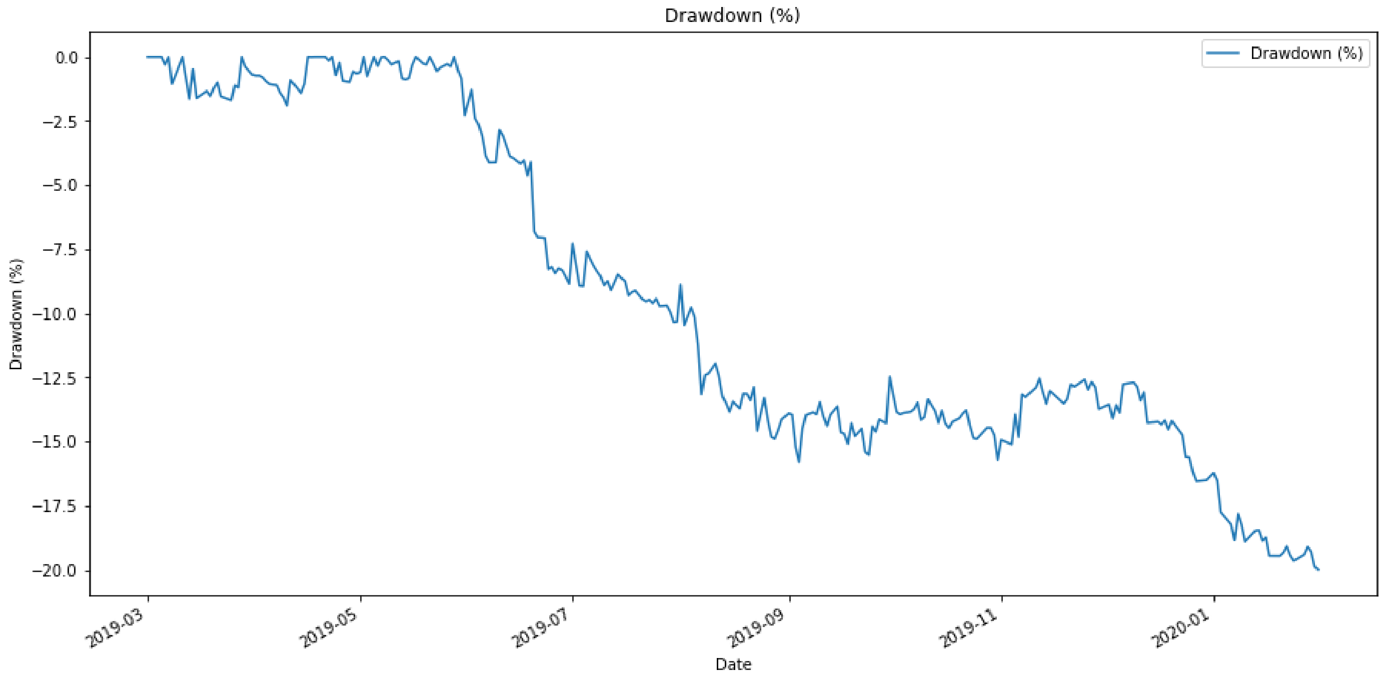
The plot below is the spread evoluation for the out-of-sample period, it is clear that there is a break in the relationship estimated from the past observations, mainly due to the looming global crisis that spiked the price of Gold as a hedging instrument opening up the gap against the Japanese yen.
This serves as a good illustration that it is critical to have a close monitoring of the validity of the relationship estimates as they are short-lived

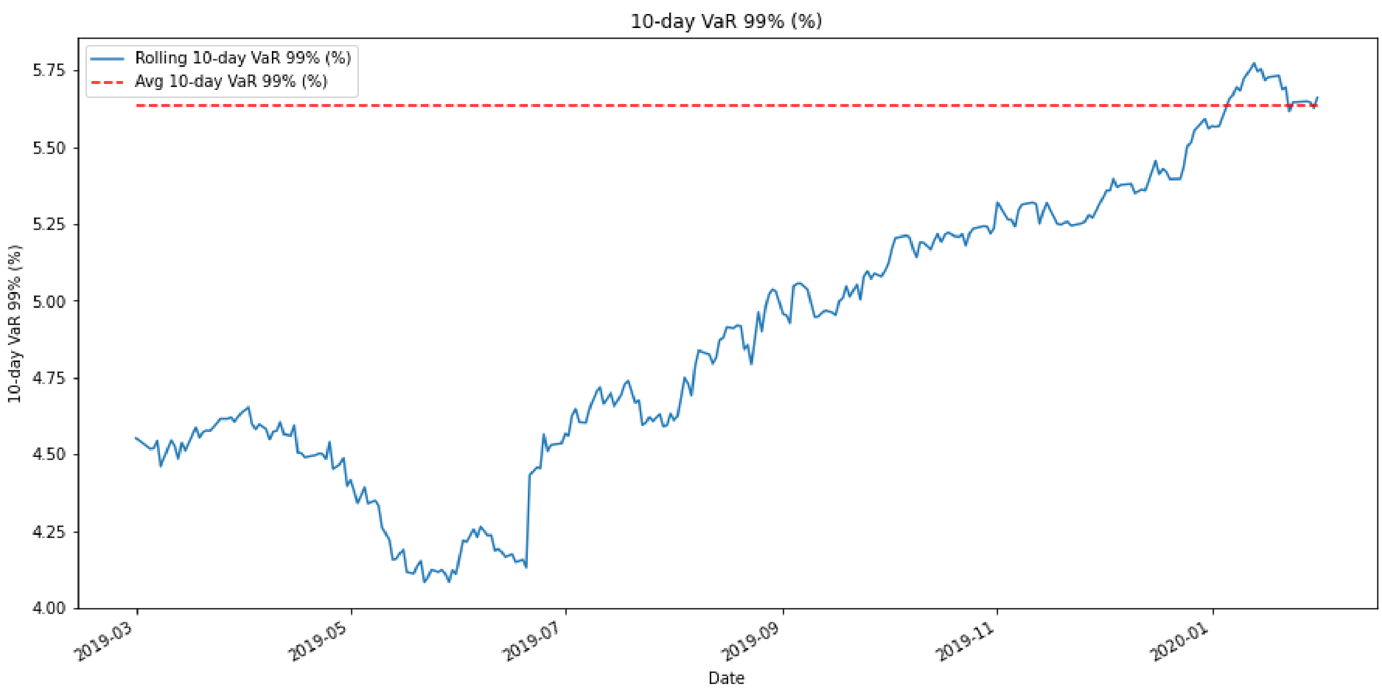

Pairs Trading JPYUSD=X x GC=F with Z-Score = 0.2
Long Run Mean μ_e
Upper/Lower Bounds
Entry Point
Exit Point
Spread
Spread
Date


Rolling Sharpe 6M
Rolling Sharpe 6M (%)
Avg Sharpe = -189.88%
Rolling Shape 6M (%)
Date


Rolling β 6M
Rolling β 6M
Avg β = 0.10
Rolling β 6M
0.2
0.1
0.0
−0.1
2019-03
2019-05
2019-07
2019-09
2019-11
2020-01
Date


Rolling Annualized α 6M
Rolling α 6M (%)
Avg α = -17.83%
Rolling Annualized α 6M (%)
10
0
−10
−20
−30
2019-03
2019-05
2019-07
2019-09
2019-11
2020-01
Date





Pair 3: Canadian Dollar (CADUSD) x Oil Futures (BZ=F)

Finally, I analyze the final Canadian Dollar x Oil Futures - the thesis behind this pair is that Canadian economy is heavily dependent on Oil prices and that any discrepancy between its currency and oil prices would eventually faded, therefore making a good candidate for long/short strategy. From previous charts we noticed that the Canadian currency and Oil prices have strong correlation (0.90) but this does not guarantee cointegration i.e. mean-reverting spread or stationary residuals

Engle Granger Fitting

I apply once again the Engle-Granger cointegration estimation to evaluate if this pair is eligible for a long/short trading strategy.

Table 20 shows the fitting of two linear regressions Canadian dollar as function of Oil prices and vice-versa (oil as function of Canadian Dollar).
Noticed that all coefficients are highly significant but this is not enough and in fact is spurious correlation.

Table 20: CADUSD and Oil OLS Regression Results

	Variable				**BZ=F**	**Variable**				**CADUSD=X**
Statistic		**Beta**	**P-Value**	**SE**	**T-Stat**		**Beta**	**P-Value**	**SE**	**T-Stat**
0	Constant	-1.720	0.000	0.031	-55.717	Constant	0.655	0.000	0.003	192.765
1	CADUSD=X	2.950	0.000	0.032	91.511	BZ=F	0.273	0.000	0.003	91.511

Table 21 shows the results for the ADF test applied to both estimate equations in order to assess if the spread between these two securities is stationary. From the table we note that for Equation 1 the ADF T-statistic is -2.38 whereas for Equation 2 is -2.16 neither is statistically at the 10% level. **Hence we do not reject the null hypothesis that the series has a unit root i.e. we do not have evidence that the series is stationary.**
Because we conclude from ADF Test that the series is not stationary **we cannot implement a long/short strategy based on pair trading since the spread is not mean-reverting and in fact has a direction.**

Table 21: CADUSD and Oil ADF Test Results

	Equation 1:	**BZ=F = CADUSD=X + Constant**		**Equation 2:**	**CADUSD=X = BZ=F + Constant**	
Null Hypothesis	Series has unit root			Series has unit root		
T-Statistic	-2.383868			-2.176502		
P-Value (MacKinnon)	0.146341			0.214884		
Optimal Lag	2			2		
Confidence Level	1%	5%	10%	1%	5%	10%
Mackinnon Critical Value	-3.433569	-2.862962	-2.567527	-3.433569	-2.862962	-2.567527
Reject/Not Reject H0	Not Reject	Not Reject	Not Reject	Not Reject	Not Reject	Not Reject
Stationary/Non Stationary	Non Stationary	Non Stationary	Non Stationary	Non Stationary	Non Stationary	Non Stationary

Table 22 is the error correction for Equation 1 and even though the residual is not stationary, the $-(1-\alpha)$ coefficients (Residual(-1) in the table) is statistically significant. But because, the process is not stationary we do not proceed to compute the OU parameters since they would be meaningless

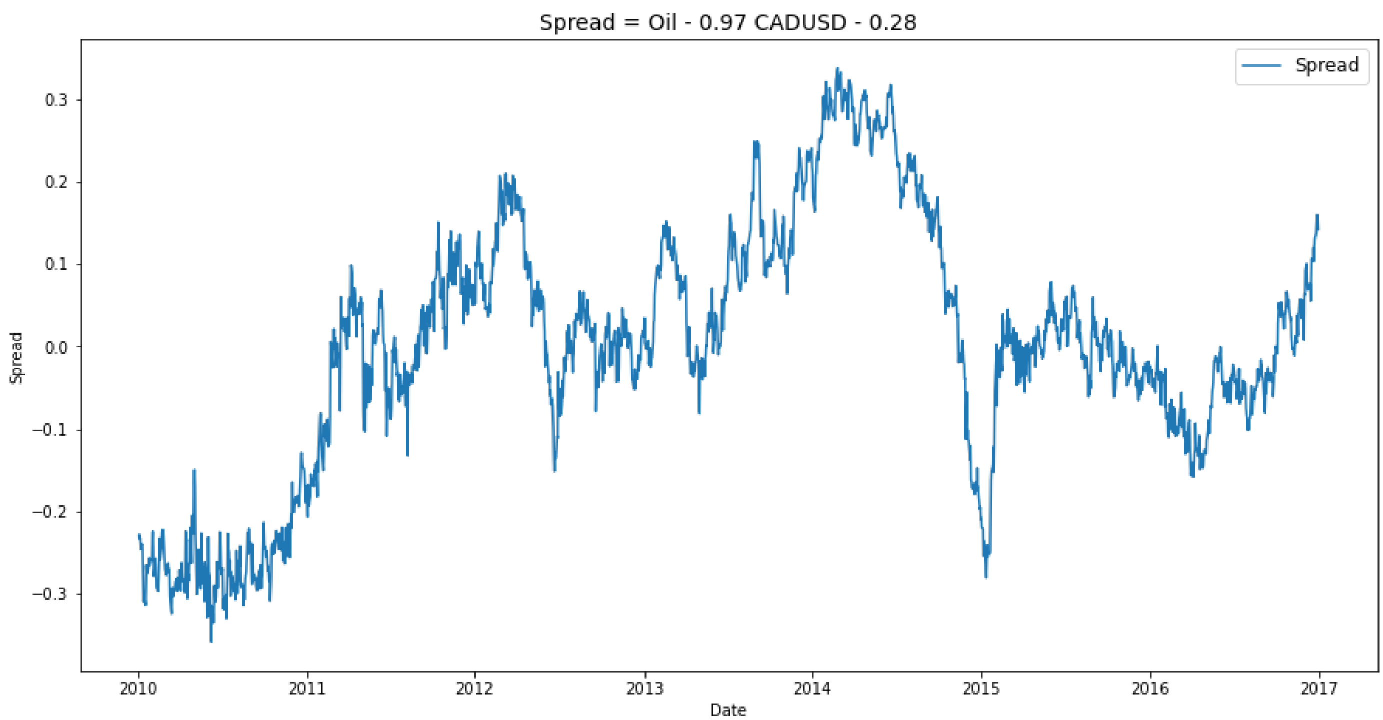
Table 22: Equation 1: Oil = CADUSD + Const
Error Correction Model Results

	dY			
Statistic	**Beta**	**P-Value**	**SE**	**T-Stat**
Variable				
dX	0.749	0.000	0.077	9.674
Residual(-1)	-0.007	0.003	0.002	-2.940

As a final note, it is intersting to illustrate the non stationary of the Canadian Dollar x Oil pair on the chart below. This spread was computed as

$$Spread_t = Oil_t - 2.95 CADUSD_t + 1.720$$

It is rather clear that this is not a stationary process as there is a hike on the spread between 2010 and 2012 followed by a less pronounced trend but with non constant variance as the ups and downs are ever more pronounced. There is a sharp decline in oil prices (hence reducing the computed spread) around 2015 due to intensified fracking on US.
This is an interesting case because the correlation is very high but there is no arbitage opportunity for pair trading.



Conclusion

Pairs trading is a trading strategy that relies on spread between two assets that show mean-reverting (Stationary) properties that can be exploited in the short-term as a sort of market-neutral strategy as it is long and short on assets that are cointegrated i.e. they share some underlying fundamental growth that any discrepancies will eventually vanish.

To properly build these strategies the application of robust statistical methodologies such as the Engle-Granger are necessary but **do not guarantee sucessful trading results** as shown in this study, due to:

1. The market is dynamic and requires constant reevaluation and close monitoring of the stability/validity of estimated parameters, no restimation and constant reassessment is a recipe for failure especially for not so strongly cointegrated pairs such as Japanese Yen x Gold
2. Essentially one should focus on economic priors and feasiblity on selecting pairs, the most successful one in this study was between RBC and TD, the two largest Canadian banks that share the economic fundamentals making a great pair trading
3. Short-term trade are much more efficient due to lower exposure to market movement, hence limiting the exposure to breakdown on the relationships estabilished.
4. High correlation does not guarantee cointegration, hence proper statistical evaluation is necessary prior to designing trading strategy

References

1. CQF Lectures Notes
2. CQF Final Project Workshop
3. Seabold, Skipper, and Josef Perktold. “statsmodels: Econometric and statistical modeling with python.” Proceedings of the 9th Python in Science Conference. 2010.


```
[NbConvertApp] Converting notebook CQF_PairsTrading.ipynb to html
[NbConvertApp] Writing 1763943 bytes to CQF_PairsTrading.html
```