Carlos Salas Najera

CQF Project: Time Series Analysis and Backtesting

June, 2019

Pair Trading and VAR analysis applied to Energy Stocks

Abstract

This project is based on developing a straightforward application to apply a simple mean-reversion strategy tailored to Energy stocks (XOM, COP, CVX) and two ETFs (XLE and SPY). Firstly, a VAR model is built using security price returns as first step before conducting IRF (Impulse Response Function) analysis and Granger Causality tests to obtain preliminary clues about the most interesting pair trading combinations. Secondly, 2-step Engle Granger is used to provide a formal framework to shortlist the best pair combinations, with the first stage based on testing each pair for cointegration, and the second step allowing to test for the existence of an ECM (Error Correction Mechanism) that brings the cointegrated pair ultimately towards a long term equilibrium. The result obtained downsized significantly the pair combinations more likely to be successful with COP-XOM as one of the best choices. The last section delves into backtesting of the COP-XOM pair starting by identifying key strategy parameters using an Ornstien-Uhlenbeck (OU) process to model the spread. Finally, the backtested strategy returns are compared against a sample of Fama-French risk factors in order to ascertain whether or not the alpha delivered by the cointegration exit/entry points is significant.

Important: the user can find more information about the files attached and instructions to load the data in the first section of the Appendix at the end of this document.

Keywords: Pairs Trading, Energy, Multi-Dimensional, VAR, Backtesting, IRF.

1. Introduction

Pairs Trading (aka Statistical Arbitrage) was first developed at Morgan Stanley during the 80s and it is based on finding relative mispriced relationships between a pair or group of stocks in order to benefit from an ultimate convergence towards its theoretical equilibrium.

The strategy has several versions relying on multiple methods with the three key branches being the distance method (DM), the cointegration method (CM), and the stochastic spread method (SP). On the one hand, DM is the simplest pair trading strategy with the largest body of research, and consists in calculating the spread between the normalized prices of those possible combinations of stock pairs that have the least sum of squared spreads. Authors such as Gatev, Goetzmann and Rouwenhorst (2006) found DM to generate positive P&L for the period 1962-2002, whereas Do and Faff (2010) pointed out a that the number of loss-making trades was high and increasing after the turn of the century, which make DM a pair trading strategy with diminishing returns as time passes due to its broad adoption and, especially, after accounting for transaction costs.

On the other hand, CM is a parametric model based in finding a pair of stocks whose linear combination is stationary i.e. both stocks are said to be cointegrated. Cointegrated pairs are therefore traded based on the stationarity of their spread and short term deviations from the long term equilibrium. Rad, Kwong Yew Low and Faff (2016) conducted a thorough analysis using US stocks' data from July 1962 to December 2014 comparing DM, CM and copula methods. Their findings highlighted DM delivering a slightly higher monthly return than CM, but CM exhibited a slightly superior Sharpe ratio. Both DM and CM perform equally well over the full sample period outperforming in terms of profits and downside risk during periods of market turmoil. Stochastic spread methods (SP), such as Ohrnstein-Uhlenbeck process, verify the existence of a well-known stochastic model for mean reversion.

Mazo, Vaquero, Gimeno (2016) apply an enhanced DM pair trading strategy applied to European equities introducing idiosyncratic risk variables that allow to anticipate persistent divergences in the pairs and reduce the number of losing trades. In this way, next 12-month Consensus Earnings per Share (EPS), Book Value Per Share (BVPS) and Consensus Target Price (TP) variables lead to an average +35% return improvement tantamount to an additional +90 bps per month or +11% annually. Jacobs and Weber (2014) find as well that US stocks pairs consisting of stocks with high idiosyncratic risk are indeed more profitable than low idiosyncratic risk ones with a statistically significant monthly difference of more than 40 bps.

Amihud (2002) is among the first to study illiquidity as one of the sources of profitability of a pair trading strategy. Later in the decade, Engelberg, Gao, Jagannathan (2009) find a significant negative relationship between the profitability of a DM pair trading strategy and the next factors: size (market capitalization), liquidity, sell-side coverage, and institutional ownership. Riedinger (2017) also conducts a DM pair trading strategy on US stocks pointing out that high pair volatility and low correlation are beneficial for the return per trade, but negative for the trading frequency, influencing dramatically the return per trade and the trading frequency

Regardless of the method used, the same questions need to be answered to implement an efficient and profitable pairs trading strategy as stated by Vidyamurthy (2004): How do we calculate the spread? How do we identify stock pairs for which such a strategy would work? When can we say that the spread has substantially diverged from the mean?.

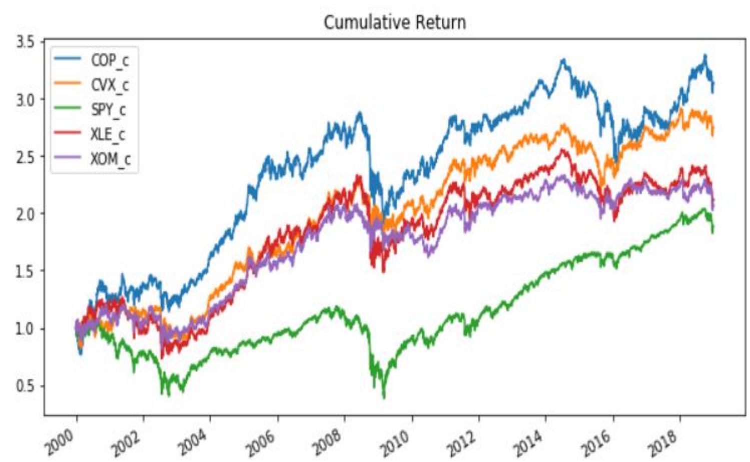
This project is based on one of the CQF suggested topics: "Time Series Analysis and Backtesting" following the first design recommended in the document "Certificate in Quantitative Finance Final Project Brief" named "Design 1: Learning and Cointegration in Pairs" covering topics such as:

1. Implement concise matrix form estimation for multivariate.
2. IRF analysis and Granger Causality test (optional)
3. Implement Engle-Granger procedure.
4. Strategy Backtesting.

2. Data

A sample of 3 stocks and 2 ETFs has been chosen for the period 2000-2018 using daily data from yahoofinance via Python's pandas_datareader library:

Ticker	Name	Weight XLE	Weight SPY
COP	ConocoPhillips	~5%	~0.3%
CVX	Chevron Corporation	~22%	~1.5%
XOM	Exxon Mobil Corporation	~22%	~1.5%
XLE	Energy Select Sector SPDR	NA	NA
SPY	SPDR S&P 500 ETF Trust	NA	NA



The three stocks belong to the Oil & Gas Integrated industry, part of the Energy Sector (MSCI GSCI Sector Classification). The frequency of the data is daily with observations ranging from 31^{st} December 1999 to 31^{st} December 2018.

An integrated oil and gas company is a business entity that engages in the exploration, production, refinement and distribution of oil and gas. Hence, an integrated company is more diversified than a pure E&P (exploration & production) firm or a firm specialized in refining crude oil to obtain other products such as petroleum naphtha, gasoline, diesel fuel, asphalt base, heating oil, kerosene, liquefied petroleum gas, jet fuel and fuel oils. The three stocks chosen are the top three oil integrated majors by market capitalization and the largest companies in the US Energy sector. In other words, these stocks have identical nature of their operations (integrated oil), similar market capitalization classification (mega large caps), look alike business fundamentals as well as high liquidity in terms of their stock trading turnover that make them ideal candidates for a pair trading strategy after the observations made about these features in section 1.

In addition, two ETFs have been added to our sample. On the one hand, Energy Select Sector SPDR seeks to provide an effective representation of the energy sector of the S&P 500 Index. On the other hand, SPDR S&P 500 ETF Trust allows investors to track the S&P 500 stock market index and is the largest ETF in the world.

3. Exploratory Data Analysis

Security returns are calculated from the original level prices and multiple statistical studies and density plots are applied to the original returns and a standardized version as displayed below. When comparing all five security returns against a simulated standard normal distribution, the results against the hypothesis of normality are staggering with all securities experiencing abnormal figures of both skewness and kurtosis compared to the automatically generated normal distribution.

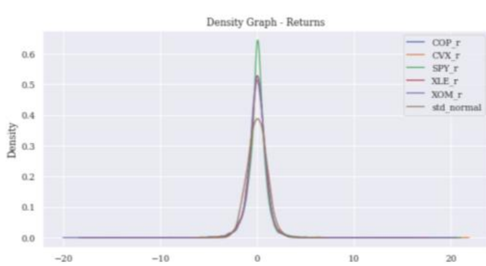
The last row shows the results of testing the null hypothesis of normality based on D'Agostino and Pearson's tests - that combine skew and kurtosis to produce an omnibus test of normality – and the ultimate results confirm our preliminary analysis conclusions: the returns cannot be said to be normally distributed.

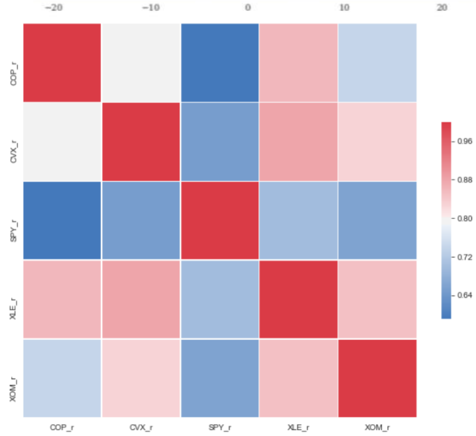
Daily Returns - Original Prices

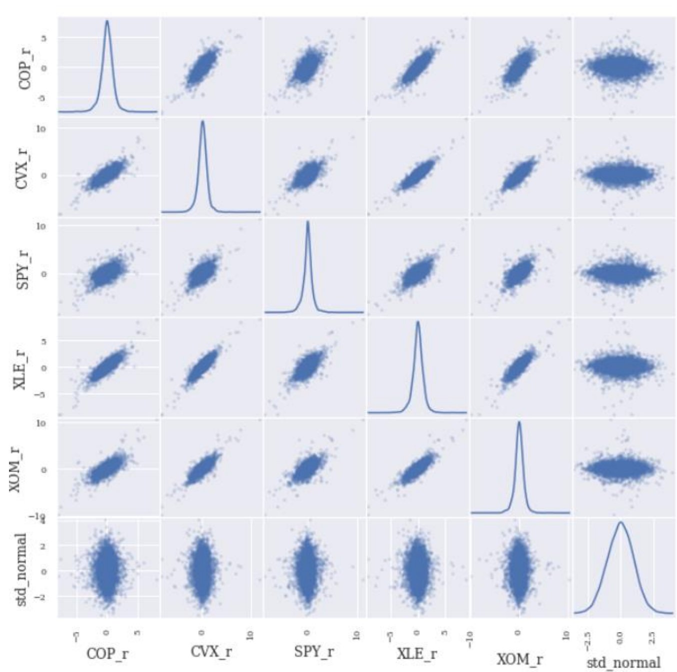
	COP_r	CVX_r	SPY_r	XLE_r	XOM_r
count	4779	4779	4779	4779	4779
mean	0.00045	0.00037	0.00019	0.00023	0.00022
std	0.01892	0.01601	0.01212	0.0172	0.0152
min	-0.14869	-0.13341	-0.10364	-0.156	-0.15027
25%	-0.00901	-0.00776	-0.00476	-0.00807	-0.00725
50%	0.00081	0.00078	0.00063	0.00063	0.00039
75%	0.01092	0.00879	0.00577	0.00934	0.00804
max	0.15365	0.18942	0.13558	0.1525	0.15863
skew	-0.309584	0.0491737	-0.0571121	-0.422637	0.0149614
kurt	5.47089	10.6404	10.4491	9.06688	10.484
normal?	No	No	No	No	No

Daily Returns - Standardized

	COP_r	CVX_r	SPY_r	XLE_r	XOM_r	std_normal
count	4779	4779	4779	4779	4779	4779
mean	0	0	-0	0	-0	-0.00308
std	1.0001	1.0001	1.0001	1.0001	1.0001	0.99811
min	-7.88418	-8.35481	-8.56654	-9.08539	-9.90165	-3.12225
25%	-0.4999	-0.50766	-0.40837	-0.48289	-0.49193	-0.67559
50%	0.01902	0.02607	0.03696	0.02317	0.01135	0.00309
75%	0.55384	0.52601	0.46089	0.52976	0.51468	0.66588
max	8.09876	11.8072	11.1713	8.85511	10.4224	3.57185
skew	-0.309584	0.0491737	-0.0571121	-0.422637	0.0149614	0.0283541
kurt	5.47089	10.6404	10.4491	9.06688	10.484	-0.0532683
normal?	No	No	No	No	No	Yes







4. Matrix form estimation for multivariate regression

This section will apply a VAR (Vector Auto Regression) model to security price daily returns. VAR(p) models assume that the passed time series are stationary. Hence, should we want to conduct a direct analysis of non-stationary time series, a standard stable VAR(p) model will not appropriate. One solution for non-stationary or trending data is to transform it into stationary by first-differencing or some other method.

Confirming Stationarity Input

In this way, ADF (Augmented Dickey Fuller) test for stationarity in the security returns confirms that the security returns to be utilized in our VAR(p) model are stationary. The tables displayed below underscore significantly small p-values rejecting the null hypothesis that a unit root is present aka non-stationarity. In addition, security price differences are also tested for unit root existence and the results also are confirming stationarity of the data. Finally, security price level data is tested as well and, as expected, the results show significant p-values that force us to do not reject the null hypothesis of existence of unit root and, as a result, our VAR(p) analysis will be based on daily price returns as it was intended at the outset.

ADF test results were yielded from a function developed in the utility_py file enclosed with this project. The function uses StatsModels 's Python library function adfuller() whereby different ADF lag-driven models are tested using AIC minimization and introducing Schwert (1989) rule of thumb:

$$p_{\max} = \left[12 \left(\frac{n}{100} \right)^{1/4} \right]$$

ADF Test Price Returns

	ADF Stat	ADF p-value	Lag	1% CV	5% CV	10% CV
COP_r	-52.6552	0.0	1	-2.5662	-1.9411	-1.6168
CVX_r	-28.9776	0.0	5	-2.5662	-1.9411	-1.6168
SPY_r	-33.3905	0.0	4	-2.5662	-1.9411	-1.6168
XLE_r	-33.4838	0.0	4	-2.5662	-1.9411	-1.6168
XOM_r	-33.9273	0.0	4	-2.5662	-1.9411	-1.6168

ADF Test Price Returns

	ADF Stat	ADF p-value	Lag	1% CV	5% CV	10% CV
COP_d	-71.9650	0.0	0	-2.5662	-1.9411	-1.6168
CVX_d	-28.6403	0.0	5	-2.5662	-1.9411	-1.6168
SPY_d	-32.7963	0.0	4	-2.5662	-1.9411	-1.6168
XLE_d	-28.7683	0.0	5	-2.5662	-1.9411	-1.6168
XOM_d	-33.4440	0.0	4	-2.5662	-1.9411	-1.6168

ADF Test Price Levels

	ADF Stat	ADF p-value	Lag	1% CV	5% CV	10% CV
COP	0.5259	0.8308	1	-2.5662	-1.9411	-1.6168
CVX	0.9291	0.9058	5	-2.5662	-1.9411	-1.6168
SPY	1.9087	0.9875	5	-2.5662	-1.9411	-1.6168
XLE	0.0707	0.7073	2	-2.5662	-1.9411	-1.6168
XOM	0.3680	0.7928	5	-2.5662	-1.9411	-1.6168

VAR Model Theoretical Framework

The matrix form for the VAR model can be expresses as it follows for an observation t:

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-1} + \cdots + A_p y_{t-p} + A_q y_{t-q} + e_t$$

Where:

- K = number of endogenous variables
- Y_t = K x 1 array
- A = K x K matrix
- e_t = K X 1 array and error term.

Writing a more comprehensive matrix notation aka large matrix notation:

$$\begin{vmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{k,t} \end{vmatrix} = \begin{vmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{vmatrix} + \begin{vmatrix} a^1_{1,1} & a^1_{1,2} & \cdots & a^1_{1,k} \\ a^1_{2,1} & a^1_{2,2} & \cdots & a^1_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ a^1_{k,1} & a^1_{k,2} & \cdots & a^1_{k,k} \end{vmatrix} \begin{vmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{k,t-1} \end{vmatrix} + \cdots + \begin{vmatrix} a^p_{1,1} & a^p_{1,2} & \cdots & a^p_{1,k} \\ a^p_{2,1} & a^p_{2,2} & \cdots & a^p_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ a^p_{k,1} & a^p_{k,2} & \cdots & a^p_{k,k} \end{vmatrix} \begin{vmatrix} y_{1,t-p} \\ y_{2,t-p} \\ \vdots \\ y_{k,t-p} \end{vmatrix} + \begin{vmatrix} e_{1,t} \\ e_{2,t} \\ \vdots \\ e_{k,t} \end{vmatrix}$$

This expression can be also rewrite using a concise matrix notation which includes T+1 observations y_0 through y_T: $Y = BZ + U$

Where:

$$Y = \begin{vmatrix} y_p & y_{p+1} & \cdots & y_T \end{vmatrix} = \begin{vmatrix} y_{1,p} & y_{1,p+1} & \cdots & y_{1,T} \\ y_{2,p} & y_{2,p+1} & \cdots & y_{2,T} \\ \vdots & \vdots & \ddots & \vdots \\ y_{k,p} & y_{k,p+1} & \cdots & y_{k,T} \end{vmatrix}$$

$$B = \begin{vmatrix} c & A_1 & A_2 & \ldots & A_p \end{vmatrix} = \begin{vmatrix} c_1 & a^1_{1,1} & a^1_{1,2} & \cdots & a^1_{1,k} & \cdots & a^p_{1,1} & a^p_{1,2} & \cdots & a^p_{1,k} \\ c_2 & a^1_{2,1} & a^1_{2,2} & \cdots & a^1_{2,k} & \cdots & a^p_{2,1} & a^p_{2,2} & \cdots & a^p_{2,k} \\ \vdots & \vdots & \ddots & \vdots & & & & & & \\ c_k & a^1_{k,1} & a^1_{k,2} & \cdots & a^1_{k,k} & \cdots & a^p_{k,1} & a^p_{k,2} & \cdots & a^p_{k,k} \end{vmatrix}$$

$$Z = \begin{vmatrix} 1 & 1 & \dots & 1 \\ y_(p-1) & y_(p) & \dots & y_(T-1) \\ y_(p-2) & y_(p-1) & \dots & y_(T-2) \\ \vdots & \vdots & \ddots & \vdots \\ y_(0) & y_(1) & \dots & y_(T-p) \end{vmatrix} = \begin{vmatrix} 1 & 1 & \dots & 1 \\ y_{1,p-1} & y_{1,p} & \cdots & y_{1,T-1} \\ y_{2,p-1} & y_{2,p} & \cdots & y_{2,T-1} \\ \vdots & \vdots & \ddots & \vdots \\ y_{k,p-1} & y_{k,p} & \cdots & y_{k,T-1} \\ y_{1,p-2} & y_{1,p-1} & \cdots & y_{1,T-2} \\ y_{2,p-2} & y_{2,p-1} & \cdots & y_{2,T-2} \\ \vdots & \vdots & \ddots & \vdots \\ y_{k,p-2} & y_{k,p-1} & \cdots & y_{k,T-2} \\ \vdots & \vdots & \ddots & \vdots \\ y_{1,0} & y_{1,1} & \cdots & y_{1,T-p} \\ y_{2,0} & y_{2,1} & \cdots & y_{2,T-p} \\ \vdots & \vdots & \ddots & \vdots \\ y_{k,0} & y_{k,1} & \cdots & y_{k,T-p} \end{vmatrix}$$

$$U = \begin{vmatrix} ep & ep+1 & \dots & e_T \end{vmatrix} = \begin{vmatrix} e_{1,p} & e_{1,p+1} & \dots & e_{1,T} \\ e_{2,p} & e_{2,p+1} & \dots & e_{2,T} \\ \vdots & \vdots & \ddots & \vdots \\ e_{k,p} & e_{k,p+1} & \dots & e_{k,T} \end{vmatrix}$$

Hence, the solution for a VAR(p) model coefficients in B matrix can be obtained easily using, for instance, OLS (ordinary least squares estimation) and the expression:

$$\hat{B} = (Z^T Z)^{-1} ZY$$

Finding Optimal p lag for VAR(p) model

After confirming stationarity of the returns calculated for our securities, it is necessary to step forward to the next step i.e. calculate the VAR model optimal number of lags (p). The methodology is straightforward consisting on testing various lags for p and select the value that minimizes a function called the "Information Criterion". Several authors have come up with different alternatives to obtain an information criterion (IC) that optimizes model selection such as Akaike Information Criterion (AIC), Schwarz-Bayes Criterion (SBC) – also known as the Bayesian

Information Criterion (BIC) – Akaike's Final Prediction Error Criterion (FPE), and Hannan-Quinn Criterion (HQ).

Each criterion is a sum of two terms, one that characterizes the entropy rate or prediction error of the model, and a second term that characterizes the number of freely estimated parameters in the model (which increases with increasing model order). By minimizing both terms, we seek to identify a model that does not overfit the data with too many parameters (parsimonious) without neglecting modelling accuracy. Four main IC are detailed below along with their mathematical expression (Lütkepohl, 2005):

Estimator	Formula
Schwarz-Bayes Criterion (Bayesian Information Criterion)	$SBC(p) = ln\left\|\tilde{\Sigma}(p)\right\| + \frac{ln(\hat{T})}{\hat{T}}pM^2$
Akaike Information Criterion	$AIC(p) = ln\left\|\tilde{\Sigma}(p)\right\| + \frac{2}{\hat{T}}pM^2$
Akaike's Final Prediction Error	$FPE(p) = \left\|\tilde{\Sigma}(p)\right\| + \left(\frac{\hat{T}+Mp+1}{\hat{T}-Mp-1}\right)^M$ and its logarithm (used in SIFT) $ln(FPE(p)) = ln\left\|\tilde{\Sigma}(p)\right\| + Mln\left(\frac{\hat{T}+Mp+1}{\hat{T}-Mp-1}\right)$
Hannan-Quinn Criterion	$HQ(p) = ln\left\|\tilde{\Sigma}(p)\right\| + \frac{2ln(ln(\hat{T}))}{\hat{T}}pM^2$

Expression	**Definition**
$\log(\|(\tilde{\sum}(p)\|$	logarithm of the determinant of the estimated noise covariance matrix (prediction error) for a VAR model of order p
$T' = T * N$	Total number of datapoints used to fit the model
$pM^2 = n * (n * j + 1)$	N being the number of securities (time series) and j=lag

Each IC is an estimate of a constant plus the relative distance between the unknown true likelihood function of the data and the fitted likelihood function of the model. Hence, for each method, the best model minimizes IC and, as a consequence, it is the one minimizing the residuals too and, thereby, the determinant of the covariance matrix will be approximately zero:

$$p_{optimal} = \arg\min_{p} IC(p)$$

The key difference between the criteria is how severely each penalizes model complexity (the second term). AIC and SBC are the most widely used, but SBC more heavily penalizes larger model orders. For moderate and large samples, FPE and AIC are essentially equivalent; however, FPE may outperform AIC for very small sample sizes. HQ penalizes high model orders more heavily than AIC but less than SBC. Both SBC and HQ are consistent estimators, yet this cannot be said of AIC and FPE.

Optimal lag selection using a single IC might be dangerous as the optimal lag for both AIC and FPE (p=15) yields a very complex model that can be prone to overfitting or only suitable for a single specific purpose (describe or forecast). When considering BIC and HQIC's optimal lags the models are much less complex. As a result, an IC score ranking is calculated (the higher the IC, the higher the score) to gather different insights from the multiple IC considered and the results' ranking displayed in the next table were favoring the use of a VAR model with 2 lags:

	aic	bic	hqic	fpe	score
2	7.0	3.0	1.0	7.0	4.50
4	8.0	5.0	5.0	8.0	6.50
10	5.0	11.0	11.0	5.0	8.00
5	10.0	6.0	6.0	10.0	8.00
9	6.0	10.0	10.0	6.0	8.00
3	13.0	4.0	3.0	13.0	8.25
15	1.0	16.0	16.0	1.0	8.50
8	9.0	9.0	9.0	9.0	9.00
16	2.0	17.0	17.0	2.0	9.50
17	3.0	18.0	18.0	3.0	10.50
1	20.0	2.0	2.0	20.0	11.00
7	14.0	8.0	8.0	14.0	11.00
18	4.0	19.0	19.0	4.0	11.50
11	11.0	12.0	12.0	11.0	11.50
0	21.0	1.0	4.0	21.0	11.75
12	12.0	13.0	13.0	12.0	12.50
6	18.0	7.0	7.0	18.0	12.50
13	16.0	14.0	14.0	16.0	15.00
14	15.0	15.0	15.0	15.0	15.00
19	17.0	20.0	20.0	17.0	18.50
20	19.0	21.0	21.0	19.0	20.00

Next step is to run the VAR model using the optimal 2 lags. For this purpose, a coded proprietary function is imported from the enclosed file py_utility.py and the results are matched against the results of a VAR (2) run using Python's well-known StatsModels library. The coefficients obtained from both the proprietary coded VAR(2) function and SM function are highlighted in the next table with identical results:

Proprietary VAR(2) Coefficients

	COP_r	CVX_r	SPY_r	XLE_r	XOM_r
const	0.000509	0.000433	0.000233	0.000291	0.000303
L1.COP_r	-0.018881	-0.018886	0.014031	0.012353	-0.033281
L1.CVX_r	-0.032208	-0.062964	-0.026003	-0.046691	-0.012941
L1.SPY_r	0.030181	-0.017600	-0.024006	0.062370	-0.076478
L1.XLE_r	0.051456	0.068997	0.002635	-0.025375	0.037449
L1.XOM_r	-0.083986	-0.056896	-0.042198	-0.045301	-0.060522
L2.COP_r	-0.057397	-0.020381	-0.006636	-0.030388	-0.028990
L2.CVX_r	-0.003384	-0.070288	-0.029997	-0.039309	-0.016712
L2.SPY_r	0.023969	0.017239	-0.067779	0.014367	0.003271
L2.XLE_r	0.021542	0.040624	0.014148	0.021008	0.004973
L2.XOM_r	-0.045915	-0.027369	0.015969	-0.053863	-0.062408

SM VAR(2) Coefficients

	COP_r	CVX_r	SPY_r	XLE_r	XOM_r
const	0.000509	0.000433	0.000233	0.000291	0.000303
L1.COP_r	-0.018881	-0.018886	0.014031	0.012353	-0.033281
L1.CVX_r	-0.032208	-0.062964	-0.026003	-0.046691	-0.012941
L1.SPY_r	0.030181	-0.017600	-0.024006	0.062370	-0.076478
L1.XLE_r	0.051456	0.068997	0.002635	-0.025375	0.037449
L1.XOM_r	-0.083986	-0.056896	-0.042198	-0.045301	-0.060522
L2.COP_r	-0.057397	-0.020381	-0.006636	-0.030388	-0.028990
L2.CVX_r	-0.003384	-0.070288	-0.029997	-0.039309	-0.016712
L2.SPY_r	0.023969	0.017239	-0.067779	0.014367	0.003271
L2.XLE_r	0.021542	0.040624	0.014148	0.021008	0.004973
L2.XOM_r	-0.045915	-0.027369	0.015969	-0.053863	-0.062408

Proprietary – SM Coeffs

	COP_r	CVX_r	SPY_r	XLE_r	XOM_r
const	-0.0	-0.0	-0.0	-0.0	-0.0
L1.COP_r	-0.0	-0.0	0.0	0.0	-0.0
L1.CVX_r	0.0	-0.0	0.0	0.0	0.0
L1.SPY_r	-0.0	-0.0	0.0	0.0	-0.0
L1.XLE_r	-0.0	0.0	-0.0	-0.0	0.0
L1.XOM_r	0.0	0.0	0.0	0.0	-0.0
L2.COP_r	-0.0	-0.0	-0.0	-0.0	0.0
L2.CVX_r	0.0	0.0	-0.0	0.0	-0.0
L2.SPY_r	0.0	0.0	-0.0	-0.0	-0.0
L2.XLE_r	0.0	-0.0	0.0	-0.0	-0.0
L2.XOM_r	-0.0	-0.0	0.0	0.0	0.0

Regarding coefficients' stability is tested by requiring eigenvalues of each relationship matrix A to be inside the unit circle (<1). A reminder of the VAR(p) model in terms of A matrices is shown next:

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-1} + \cdots + A_p y_{t-p} + A_q y_{t-q} + e_t$$

Hence, the VAR system satisfies stability condition if and only if:

$$|\lambda I - A| = 0$$

Using the VAR (2) model via SM library the results are positive confirming the condition above of coefficient stability in our VAR(2) model.

The full VAR (2) model results are included in the Appendix, yet the correlation matrix of the residuals and the ADF test for stationarity are provided in the two tables below. Overall, ADF test conclusions reject the existence of unit root when using even conservative significance levels such as 1%. Moreover, VAR (2) model correlation between residuals is reasonable high providing a highly likelihood that a cointegrated relationship between pairs might be found.

Correlation matrix of residuals

	COP_r	CVX_r	SPY_r	XLE_r	XOM_r
COP_r	1.000000	0.790282	0.589466	0.855394	0.744536
CVX_r	0.790282	1.000000	0.652165	0.878172	0.829195
SPY_r	0.589466	0.652165	1.000000	0.702569	0.659155
XLE_r	0.855394	0.878172	0.702569	1.000000	0.853645
XOM_r	0.744536	0.829195	0.659155	0.853645	1.000000

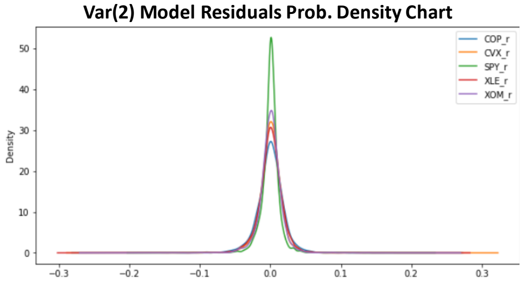
	ADF Stat	ADF p-value	Lag	1% CV	5% CV	10% CV
COP_r	-69.0456	0.0	0	-2.5662	-1.9411	-1.6168
CVX_r	-28.1499	0.0	5	-2.5662	-1.9411	-1.6168
SPY_r	-69.0690	0.0	0	-2.5662	-1.9411	-1.6168
XLE_r	-32.4648	0.0	4	-2.5662	-1.9411	-1.6168
XOM_r	-32.3506	0.0	4	-2.5662	-1.9411	-1.6168

Nevertheless, the model residuals seem to suffer from both high kurtosis and higher-than-average skew than the normal distribution making it poor model to be used for a purpose other than gaining explanatory power about the features. The following Q-Q plots demonstrate how fat tails are present in our model residuals:

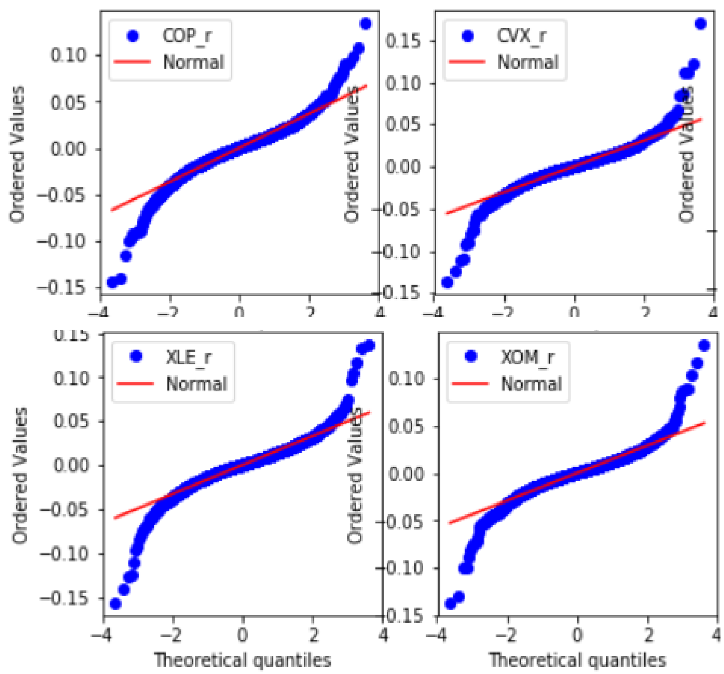
Var(2) Model Residuals Key Statistics

	COP_r	CVX_r	SPY_r	XLE_r	XOM_r
count	4777	4777	4777	4777	4777
mean	0	0	0	0	0
std	0.01885	0.01593	0.01204	0.0171	0.01503
min	-0.14317	-0.13617	-0.09408	-0.15604	-0.13579
25%	-0.0096	-0.00791	-0.00501	-0.00826	-0.00744
50%	0.00053	0.00049	0.00059	0.00047	0.00041
75%	0.01046	0.00848	0.00558	0.00916	0.00791
max	0.1342	0.16997	0.12447	0.13681	0.13473
skew	-0.346195	-0.10574	-0.205981	-0.465245	-0.250206
kurt	4.92735	9.10851	9.03467	7.92495	8.41674
normal?	No	No	No	No	No

Var(2) Model Residuals Prob. Density Chart



Var(2) Model Residuals – QQ Plot



5. Forecasting capability of regression with IRF and Granger Causality

IRF (Impulse Response Function)

Impulse responses are the estimated responses to a unit impulse in one of the variables. They are computed in practice using the MA(∞) representation of the VAR(p) process:

$$Y_t = \eta + \sum_{n=0}^{\infty} \phi_i * u_{t-i}$$

Or using a matrix form with two variables y_t and z_t:

$$\begin{vmatrix} y_t \\ y_t \end{vmatrix} = \begin{vmatrix} c_y \\ c_z \end{vmatrix} + \begin{vmatrix} a_{y,y} & a_{y,z} \\ a_{z,y} & a_{z,z} \end{vmatrix} \begin{vmatrix} y_{t-1} \\ z_{t-1} \end{vmatrix} + \begin{vmatrix} e_{y,t} \\ e_{z,t} \end{vmatrix}$$

The MA(∞) representation of this matrix form VAR process is:

$$\begin{vmatrix} y_t \\ y_t \end{vmatrix} = \begin{vmatrix} \bar{y} \\ \bar{z} \end{vmatrix} + \sum_{i=0}^{\infty} \begin{vmatrix} a_{y,y} & a_{y,z} \\ a_{z,y} & a_{z,z} \end{vmatrix} \begin{vmatrix} e_{y,t-i} \\ e_{z,t-i} \end{vmatrix}$$

The MA((moving-average) representation is important because it allows examining the interaction between the different dependent variables e.g. y_t and z_t sequences. The coefficients (e.g $a_{y,y}$. $a_{y.z}$. etc) are used to generate the effects of shocks during the stated number of periods.

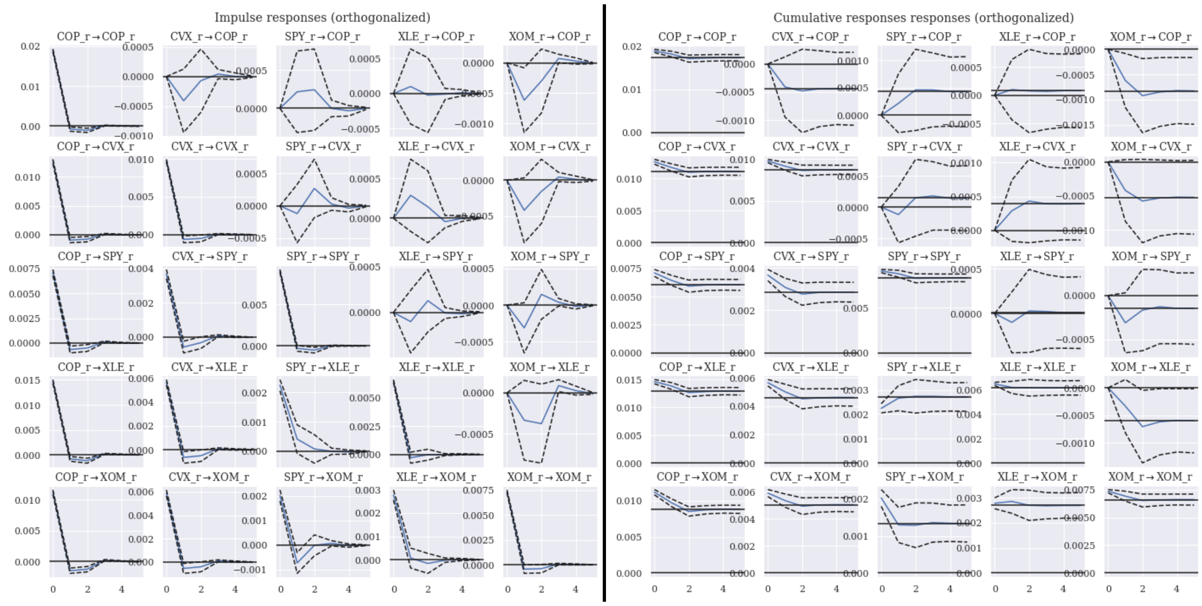
Our VAR(2) exhibited significantly high correlation between residuals. In other words, IRF assumption that shocks occur isolatedly i.e. only one variable experiences a shock on a period - is not realistic according to our results. In fact, financial markets are featured by tail-risk events during which variables tend to break historical correlation rules of thumb and misbehave increasing their correlations punctually. In this scenarios IRF analysis using MA expression is very useful as an orthogonalised transformation is carried out to obtain IRF coefficients that contain the aforementioned correlation effect. VAR model therefore is transformed into the next expression:

$$Y_t = \Theta_0 w_t + \Theta_1 w_{t-1} + +\Theta_2 w_{t-2}$$

Where:

$\Omega = PP^{-1}$ is a Cholesky decomposition
P is a lower triangular matrix
$\Theta_i = \Psi_i P$
$w_t = P^{-1} e_t$
$E(w_t w_t^{-1}) = I_N$

The results for orthogonalized IRF are shown below and point out a significant fade away effect after the initial impulse across the board for both marginal and cumulative shocks for a period of 5 days:



Granger Causality Analysis

Granger causality test was firstly proposed in 1969 and is a statistical test for studying causality between two different variables. In this way, a "causing" series is said to Granger-cause a "caused" time series if it can be shown, usually through a series of t-tests and F-tests on lagged values of the "causing" (and with lagged values of the "caused" also included), that those "causing" values provide statistically significant information about future values of the "caused" time series. The null hypothesis is that the coefficients corresponding to past values of the “causing” variable are zero i.e. the null hypothesis for the equation below - p and q being the shortest and longest significant lag lengths for x - will be:

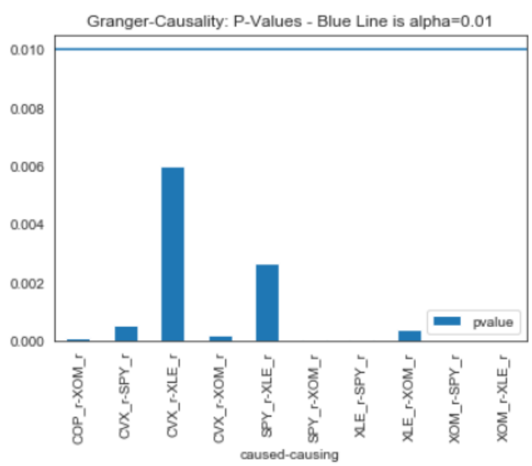
$$Ho : x_{t-p} = \ldots = x_{t-q} = 0$$

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + a_m y_{t-m} + b_p x_{t-p} + \ldots + + b_q x_{t-q} + error_t$$

Several research papers have found that the Granger Causality test has three main pitfalls. Firstly, empirical studies have found a variable "Granger-cause" another variable when both variables are non-fundamentally related e.g. increase of frogs in a lake in the US and Coca Cola Sales. Secondly, Granger causality does not account for latent confounding effects i.e. a third variable affecting both "causing" and "caused" variables making Granger Causality test results spurious. Lastly, Granger causality test does not capture instantaneous and non-linear causal relationships.

Granger causality test works under the assumption that all the VAR variables are stationary (earlier in this project price returns stationarity was confirmed). For our purposes, Granger Causality test may be a good proxy to identify preliminary significant pair directions as our fundamental analysis to test pairs (stocks within the O&G Integrated industry) and Engle-Granger test procedure will tackle most of the disadvantages aforementioned. Hence, Granger Causality tests are conducted for all the pair combinations possible for all our five securities with only the next pairs surviving for a significance level of 10% i.e. only pair combinations where p-value is below 10% are shortlisted or, in other words, only those were the null hypothesis of non-Granger Causality is rejected are considered:

caused-causing	caused	causing	pvalue
COP_r-XOM_r	COP_r	XOM_r	0.00010
CVX_r-SPY_r	CVX_r	SPY_r	0.00053
CVX_r-XLE_r	CVX_r	XLE_r	0.00602
CVX_r-XOM_r	CVX_r	XOM_r	0.00021
SPY_r-XLE_r	SPY_r	XLE_r	0.00267
SPY_r-XOM_r	SPY_r	XOM_r	0.00005
XLE_r-SPY_r	XLE_r	SPY_r	0.00004
XLE_r-XOM_r	XLE_r	XOM_r	0.00040
XOM_r-SPY_r	XOM_r	SPY_r	0.00002
XOM_r-XLE_r	XOM_r	XLE_r	0.00003



6. Engle-Granger procedure and cointegrated pair identification.

Two or more variables are said to be cointegrated when there is a linear combination of them that has a lower order of integration, then the series are said to be cointegrated and the residual of the linear combination regression is stationary i.e. two or more series are cointegrated if a linear combination exists which makes the residuals integrated of order zero I(0).

Cointegration testing is conducted in practice using three main methods: Engle–Granger two-step method, Johansen method (JM) and Phillips–Ouliaris (PO) test. As suggested by CQF project guidelines, Engle–Granger (EG) will be here to test a bivariate cointegration relationship. This method is based on testing that the residuals from the regression model between the securities y and x shown below are stationary:

$$y_t - \beta x_t = u_t$$

As mentioned before, EG will be used in this project in order to ascertain whether or not two securities in our sample are cointegrated:

Step 1: Obtain fitted residuals to conduct ADF test (Stationarity Test)

Step 2: ECM using residuals from Step 1

The training sample to be considered will be the first two thirds of data from the total period 2000-2018 i.e. sample ranging from 31st December 1999 to 30th April 2013. Moreover, a 5% significance level is used to filter p-values obtained from ADF tests in order to shortlist relevant pair cointegration combinations.

Step 1: Obtain fitted residuals to conduct ADF test (Stationarity Test)

A coded function **coint_test_bulk()** is imported from **py_utility.py** to analyze cointegration as required by CQF project guidelines. Moreover, StatsModels's library statistic is included in order to validate and double-check results obtained from the proprietary function:

	y	x	diff	ADF Lags	ADF stat	ADF stat sm	1%CV	5%CV	10%CV	ADF p-value	Cointegrated
(COP, SPY)	COP	SPY	0	5.0	-2.8	-2.8	-2.566	-1.941	-1.617	0.005	True
(COP, XLE)	COP	XLE	0	4.0	-3.2	-3.2	-2.566	-1.941	-1.617	0.001	True
(COP, XOM)	COP	XOM	0	4.0	-2.9	-2.9	-2.566	-1.941	-1.617	0.004	True
(XLE, COP)	XLE	COP	0	4.0	-3.2	-3.2	-2.566	-1.941	-1.617	0.001	True
(XLE, XOM)	XLE	XOM	0	5.0	-3.0	-3.0	-2.566	-1.941	-1.617	0.003	True
(XOM, COP)	XOM	COP	0	4.0	-2.9	-2.9	-2.566	-1.941	-1.617	0.004	True
(XOM, XLE)	XOM	XLE	0	5.0	-2.9	-2.9	-2.566	-1.941	-1.617	0.003	True

The combinations between XLE, XOM and COP price levels exhibit stationarity passing the ADF acid test with no unit root in their residuals assuming 1% significance level. The other pairs involving SPY and CVX need first differences to pass Engle Granger test with the exception of COP-SPY. XOM-COP, XLE-XOM and XOM-XLE are particularly relevant as these pair combinations were also relevant in our preliminary IRF and Granger Causality Analysis.

Step 2: ECM using residuals from Step 1

After confirming the pairs and degree of differences where we found cointegrated relationships i.e. residuals were stationary, now it's time to use the residuals to obtain the ECM (Equilibrium Correction Form). For this purpose, the pair COP-XOM has been selected due to its high statistical significance as aforementioned.

The next equation is a static equilibrium model encapsulated within the relationship between non-stationary (integrated) x and y with tau as a constant.

$$y_t = \tau_e + \beta_e x_t + e_t$$

However, using a dynamic regression model:

$$y_t = \alpha y_{t-1} + \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + e_t$$

Adding and subtracting both y_{t-1} and $b_1 x_{t-1}$:

$$y_t - y_{t-1} = \alpha y_{t-1} - y_{t-1} + \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + \beta_1 x_{t-1} - \beta_1 x_{t-1}$$

Rearranging terms the next ECM (Equilibrium Correction Model) is obtained:

$$\Delta Y_t = \beta_1 \Delta X_t - (1-\alpha) e_{t-1} + \epsilon_t$$

Where e_{t-1} are the residuals from the static equilibrium model lagged one period. This second step is helpful to understand if there's an equilibrium correction model (ECM) effectively working i.e. the parameter –(1-α) is interpreted as the speed of correction towards the equilibrium level. In fact, this parameter significance will also aid to select which pair order is more relevant. For instance, for the combination of COP and XOM the next two ECM regressions are provided to check whether COP-XOM or XOM-COP is the most relevant pair.

Dep. Variable:	COP	**R-squared:**	0.633
Model:	OLS	**Adj. R-squared:**	0.632
Method:	Least Squares	**F-statistic:**	2882.
Date:	Thu, 20 Jun 2019	**Prob (F-statistic):**	0.00
Time:	18:04:20	**Log-Likelihood:**	-922.55
No. Observations:	3351	**AIC:**	1851.
Df Residuals:	3348	**BIC:**	1869.
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.0046	0.006	0.843	0.399	-0.006	0.015
XOM	0.5560	0.007	75.924	0.000	0.542	0.570
e_t_1	-0.0038	0.002	-2.455	0.014	-0.007	-0.001

Omnibus:	368.833	**Durbin-Watson:**	1.933
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	2604.260
Skew:	-0.263	**Prob(JB):**	0.00
Kurtosis:	7.287	**Cond. No.**	4.68

Dep. Variable:	XOM	**R-squared:**	0.633
Model:	OLS	**Adj. R-squared:**	0.633
Method:	Least Squares	**F-statistic:**	2887.
Date:	Thu, 20 Jun 2019	**Prob (F-statistic):**	0.00
Time:	18:04:23	**Log-Likelihood:**	-2122.0
No. Observations:	3351	**AIC:**	4250.
Df Residuals:	3348	**BIC:**	4268.
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	3.966e-05	0.008	0.005	0.996	-0.015	0.015
COP	1.1376	0.015	75.918	0.000	1.108	1.167
e_t_1	-0.0054	0.002	-3.030	0.002	-0.009	-0.002

Omnibus:	560.150	**Durbin-Watson:**	2.073
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	6027.448
Skew:	0.455	**Prob(JB):**	0.00
Kurtosis:	9.507	**Cond. No.**	8.36

Testing the null hypothesis: −(1−α)=0, the lagged residual coefficient p-value is not significant for significance levels below 2% for both pairs. In other words, we can assume that there's a correction mechanism helping the spread to reach long term equilibrium with a level of confidence of 97%. However, XOM-COP pair equilibrium correction mechanism seems to be more reliable than COP-XOM with a p-value even below 1%.

7. Strategy Backtesting

A proprietary-coded Python class and some add-on helper functions were built and imported from the files **VecBacktester_spread.py** and **Perf_Stats_Tools.py** for backtesting and performance analysis purposes of several pair trading strategies for the selected pair XOM-COP.

In this way, Engle-Granger two-step procedure is necessary to ensure cointegration of the pair series along with stationarity of the spread:

$$\epsilon_t = y_t - \tau_e - \beta_e x_t +$$

In order to find out the optimal entry and exit points, this spread or residual of the cointegration relationship needs to be modelled. In this project a Ornstien-Uhlenbeck (OU) process is used to model the spread and identify entry/ exit points. In this way, after checking for stationarity, the spread OU process has the next solution for its SDE (stochastic differential equation) as displayed in the snapshot below from the python notebook attached in this project:

$$de_t = -\theta(e_t - \mu_e)dt + \sigma dW_t$$

Where:

- speed of reversion:

$$\theta$$

- equilibrium level:

$$\mu$$

- scatter of diffusion:

$$\sigma$$

- first difference of Wiener Process:

$$dW_t$$

OU SDE solution:

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

Transforming some terms:

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

Then it's possible to run the AR(1) regression:

$$e_t = C + Be_{t-1} + \epsilon_{t,\tau}$$

To obtain C and B coefficients that allows us to calculate:

$$\theta = -\frac{ln(B)}{\tau}$$

$$\mu_e = \frac{C}{1-B}$$

$$\sigma_{OU} = \sqrt{\frac{2\theta Var(\epsilon_{t,\tau})}{1 - e^{-2\theta\tau}}}$$

$$\sigma_{eq} \sim \frac{\sigma_{OU}}{\sqrt{2\theta}}$$

Then we can obtain the parameters required for our backtesting:

- Entry/Exit points:

$$\mu_e +/- \sigma_{eq}$$

- Approximated Speed of mean reversion:

$$\tau \sim \frac{ln(2)}{\theta}$$

Using the mean and standard deviation from the fitted OU process to define position entry/exit thresholds, a normalized z-spread can be defined as: $\frac{\hat{e}_t - \mu_e}{\sigma_{eq}}$

The parameters expressions obtained in the former snapshots are used along with all our data to build a backtesting using 3 periods: in-sample period, out-of-sample period and total period. The most useful periods to watch are the out-of-sample as it generalizes the findings from in-sample period parameter optimization as well the "total period" backtesting as it provides a lengthy time span to test the effectiveness of the strategy.

Period	**Parameter Optimization**	**Backtesting**
In-sample	31/12/1999 – 30/04/2013	31/12/1999 – 30/04/2013
Out-of-sample	31/12/1999 – 30/04/2013	01/05/2013 – 31/12/2018
Total	31/12/1999 – 30/04/2013	31/12/1999 – 31/12/2018

The backtesting trading rule rationale is summarized in the next lines using pseudo code format. Bottom line, the strategy is mean-reverting using the pair spread compared to our entry/exit points - obtained from in-sample analysis – to go long or short and, eventually, close the position back to neutral when a reversion to the mean occurs:

If position = 0:

If spread > $\mu_e + sd * \sigma_{eq}$:

Go short spread (position = 1)

elif spread < $\mu_e - sd * \sigma_{eq}$:

Go long spread (position = +1)

Else:

Keep neutral position (position = 0).

Elif current position = -1:

If spread < μ_e:

Cover Short (position = 0).

Else:

Keep Short (position = -1)

Else: (current position = +1)

If spread > μ_e:

Close Long (position = 0).

Else:

Keep Long (position = +1)

Note: sd =1 and slippage =1 day are the parameters assumed to our baseline backtesting exercise. Later sessions will be based on optimizing these two parameters.

Backtesting Results

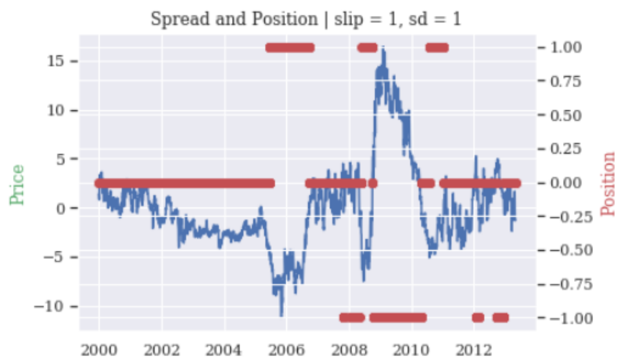
Several default assumptions are introduced in order to set forth a more realistic framework to interpret the results:

- The multiplier **sd** applied to $\boldsymbol{\sigma_{eq}}$ to obtain entry points is assumed to be equal to 1.
- The time between trading signal and trade execution is 1 day aka slippage period.
- Risk free rate assumption for Sharpe ratio calculations is 2%.
- Transaction fee rate is set up to 17 bps in line with other body of research.

The results from applying the strategy to these three different backtesting periods are shown in the next tables and charts:

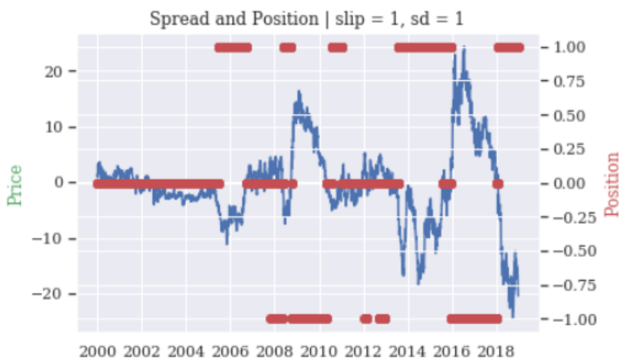
	In-sample Period			Out-of-sample Period			Total Period		
Statistics	Buy&Hold	Strategy	Benchmark	Buy&Hold	Strategy	Benchmark	Buy&Hold	Strategy	Benchmark
Gross_P&L	-69.149	75.5757	258.515	-28.0717	0.410842	-14.7601	-77.8094	76.297	205.598
Net_P&L	-69.149	70.2844	258.515	-28.0717	-0.776966	-14.7601	-77.8094	68.9614	205.598
Gross_CAGR	-8.43836	4.31002	10.0444	-5.64446	0.0723208	-2.77669	-7.61262	3.02697	6.05126
Net_CAGR	-8.43836	4.07102	10.0444	-5.64446	-0.137442	-2.77669	-7.61262	2.79694	6.05126
Net_Anual_Vol	27.9294	17.8633	35.8022	32.0298	31.0157	24.5485	29.2149	22.6092	32.8494
%_Positive	47.2	16.56	52.3	49.16	43.56	50	47.78	24.62	51.61
Skew	0.376387	-0.350431	-0.450842	0.15933	-0.0481331	-0.197	0.294759	-0.169296	-0.422532
Kurtosis	3.85613	12.7004	8.72357	5.51833	6.41453	2.14086	4.69188	11.2509	9.06787
Kurtosis PV	7.0673e-68	6.95384e-149	2.49846e-121	2.16854e-40	8.69362e-45	4.94747e-18	6.90399e-112	2.42352e-197	5.24133e-175
Downside_Vol	17.9791	23.0235	27.5964	22.5902	22.5586	17.6085	19.4754	22.7726	25.1871
Worst_Net	-0.073	-0.073	-0.156	-0.0951	-0.1248	-0.0664	-0.0951	-0.1248	-0.156
Sharpe_Ratio	-0.302848	0.226779	0.279995	-0.17685	-0.00507619	-0.113925	-0.261258	0.122824	0.183603
Sortino_Ratio	-0.470456	0.175951	0.363252	-0.250748	-0.00697923	-0.158826	-0.391911	0.121942	0.239458
Information_Ratio	-0.345701	-0.156652	None	-0.0573971	0.0692654	None	-0.260532	-0.0853606	None
Max_Drawdown	77.528	43.8476	57.3482	63.2709	39.9444	46.7568	83.8314	43.8476	57.3482
Worst_3_drawdown_avg	77.4995	43.7632	57.1828	62.866	39.703	45.8036	83.6532	43.7632	57.1828
Max_DD_Duration	12 days, 0:00:00	583 days, 0:00:00	1155 days, 0:00:00	763 days, 0:00:00	642 days, 0:00:00	87 days, 0:00:00	12 days, 0:00:00	642 days, 0:00:00	1826 days, 0:00:00

In-sample Period – Positioning and Cumulative Performance



Out-of-sample Period – Positioning and Cumulative Performance

Total Period – Positioning and Cumulative Performance



Optimization

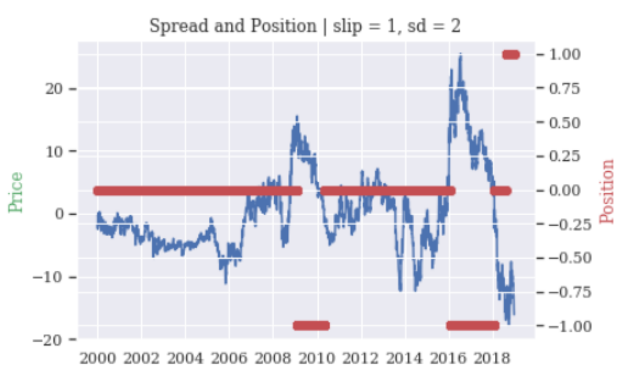
Moreover, an optimization exercise is applied to obtain both the spread's multiple applied to σ_{eq} (sd in the pseudo code presented earlier) and slippage parameters that maximizes the strategy performance. The optimization splits the total period from 2000 to 2018 in three equal subperiods and calculates key metrics multiple combinations of sd (1 to 5 with 0.5x steps) and slippage (1 to 3 days considered) returning the best parameter combination for the strategy on each subperiod.

In this way, the results exhibited below compare the total period with the optimization subperiod results with the optimal strategy being the one using a multiplier of 2x and slippage equal to 1 day. When obtaining the performance metrics for this optimal parameter combination the returns are in positive territory outperforming the "buy & hold" strategy, although falling short of the benchmark (XLE). On the bright side, the strategy is able to outperform both "buy & hold" and the benchmark in terms of risk (volatility), risk-adjusted return (Sharpe ratio) and extreme events (maximum drawdown).

Period	sd	Slip	P&L Cum	Annual Vol %	Sharpe Ratio
31/12/1999 – 12/31/2018	2	1	1.3196	13.79	0.33x
31/12/1999 – 03/05/2006	1	1	1.1396	17.89	0.23x
04/05/2006 – 28/08/2012	2	1	1.3196	13.79	0.33x
29/08/2012 – 31/12/2018	3	1	0.7439	10.29	0.29x

Optimal Strategy Performance Metrics - sd = 2 and slippage = 1 day

Statistics	Buy&Hold	Strategy	Benchmark
Gross_P&L	-70.1116	133.941	205.598
Net_P&L	-70.1116	131.961	205.598
Gross_CAGR	-6.15422	4.57134	6.05126
Net_CAGR	-6.15422	4.5246	6.05126
Net_Anual_Vol	25.6754	13.7876	32.8494
%_Positive	47.68	9.75	51.61
Skew	0.281205	0.334777	-0.422532
Kurtosis	4.81181	35.9062	9.06787
Kurtosis PV	4.21281e-114	0	5.24133e-175
Downside_Vol	17.1861	20.6222	25.1871
Worst_Net	-0.0823	-0.1123	-0.156
Sharpe_Ratio	-0.240472	0.326715	0.183603
Sortino_Ratio	-0.359257	0.218435	0.239458
Information_Ratio	-0.255959	-0.0451945	None
Max_Drawdown	76.8181	22.1121	57.3482
Worst_3_drawdown_avg	76.5893	21.8152	57.1828



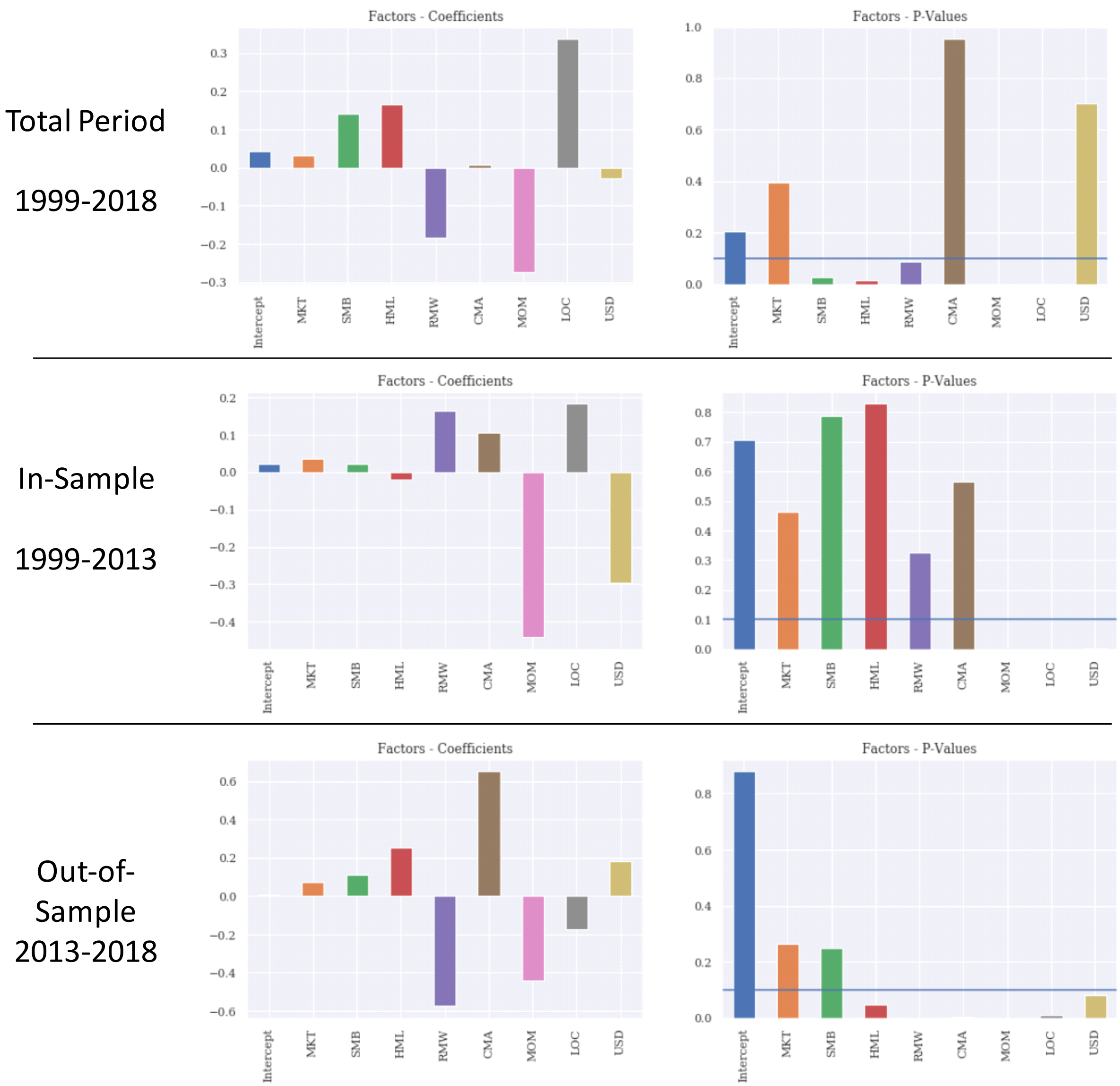
Risk Factors Exposure

This section presents a risk factor performance attribution analysis using Fama-French 2014 5-factor model comprehending Market (Market-Risk Premium), SMB (Size Premium), HML (Value Premium), RMW (Operating Profitability Premium) and CMA (Low Capital Expenditure Intensity Premium).

Moreover, Fama-French Momentum Factor has been also added to the model as several practitioners and scholars have criticized the no inclusion of the long-time accepted momentum effect in their overall 5-Factor model. Lastly, two additional factors are included to represent macro risk (USD Risk Premium) and industry/sector premium (Energy Sector Risk Premium).

Important to point out that only periods with either long or short positions in the XOM-COP spread are considered, ignoring periods where the strategy is in neutral mode. In the same fashion as it was done in past sections, we conducted risk factor analysis for three periods: total period (2000-2018) as well as our in-sample (2000-2013) and out-of-sample (2013-2018).

Several remarks can be extracted from the analysis in the charts below. Firstly, only a group of risk premium are persistent throughout time witnessing their p-values below 10% significance level: momentum (MOM) and Sector risk premium (LOC). From these couple of factors, only momentum’s factor sign is unchangeable in negative territory across all the periods. In other words, XOM-COP trading strategy is negatively exposed to periods where outperformance in momentum style prevails.



Other significant risk exposures arise when paying attention to the total period returns from 1999 to 2018. Apparently, the strategy has a positive exposure to traditional size (SMB) and value (HML) risk premiums whereas it's negatively exposed to quality factors such as Operating Profitability Premium (RMW). That said, these exposures are less persistent in time that the previously highlighted negative sensitivity to momentum (MOM).

Appendix

A1. Information about the files included

This project has been coded in Python. All the the pricing data was downloaded using pandas_datareader library into the csv file called “cqf_data_df.csv”, with the exception of Fama-French data which is automatically downloaded using pandas_datareader as part of one of the proprietary functions of one of the python files explained below. a discount brokerage in India.

The next attached files are briefly explained below:

• **CQF_Project_Notebook.ipynb**: Jupyter notebook containing an executable version of the written report that automatically loads data from different custom Python files. The notebook contains instructions to load data and the user can also open de file session_environment.txt with comprehensive information about the libraries' version utilized.

• **py_utility.py**: this file contains utility functions to carry out automatically several functions useful to proceed to load pricing data, conduct key statistics analysis, run proprietary VAR model, perform ADF test and cointegration tests using multiple securities as input at once.

• **VecBacktester_spread.py**: this file contains a Python class used to perform backtesting of a pair trading strategy. The class includes methods to obtain performance key metrics as well as to optimize multiple parameters to enhance the strategy results.

• **Perf_Stats_Tools.py**: this file contains a Python class used by Veckbacktester_spread.py. to import custom functions designed to calculate multiple strategy key performance indicators such as annualized return, volatility or Sharpe Ratio.

A2. VAR (2) Model Details

Summary of Regression Results

Model:	VAR
Method:	OLS
Date:	Thu, 20, Jun, 2019
Time:	15:56:19

No. of Equations:	5.00000	BIC:	-46.4369
Nobs:	4777.00	HQIC:	-46.4852
Log likelihood:	77256.1	FPE:	6.31473e-21
AIC:	-46.5114	Det(Omega_mle):	6.24252e-21

Results for equation COP_r

	coefficient	std. error	t-stat	prob
const	0.000509	0.000273	1.861	0.063
L1.COP_r	-0.018881	0.028326	-0.667	0.505
L1.CVX_r	-0.032208	0.038201	-0.843	0.399
L1.SPY_r	0.030181	0.032224	0.937	0.349
L1.XLE_r	0.051456	0.045196	1.139	0.255
L1.XOM_r	-0.083986	0.037033	-2.268	0.023
L2.COP_r	-0.057397	0.028348	-2.025	0.043
L2.CVX_r	-0.003384	0.038146	-0.089	0.929
L2.SPY_r	0.023969	0.032521	0.737	0.461
L2.XLE_r	0.021542	0.044890	0.480	0.631
L2.XOM_r	-0.045915	0.036654	-1.253	0.210

Results for equation CVX_r

	coefficient	std. error	t-stat	prob
const	0.000433	0.000231	1.874	0.061
L1.COP_r	-0.018886	0.023942	-0.789	0.430
L1.CVX_r	-0.062964	0.032290	-1.950	0.051
L1.SPY_r	-0.017600	0.027237	-0.646	0.518
L1.XLE_r	0.068997	0.038202	1.806	0.071
L1.XOM_r	-0.056896	0.031303	-1.818	0.069
L2.COP_r	-0.020381	0.023961	-0.851	0.395
L2.CVX_r	-0.070288	0.032243	-2.180	0.029
L2.SPY_r	0.017239	0.027489	0.627	0.531
L2.XLE_r	0.040624	0.037944	1.071	0.284
L2.XOM_r	-0.027369	0.030982	-0.883	0.377

Results for equation SPY_r

	coefficient	std. error	t-stat	prob
const	0.000233	0.000175	1.332	0.183
L1.COP_r	0.014031	0.018095	0.775	0.438
L1.CVX_r	-0.026003	0.024404	-1.066	0.287
L1.SPY_r	-0.024006	0.020585	-1.166	0.244
L1.XLE_r	0.002635	0.028872	0.091	0.927
L1.XOM_r	-0.042198	0.023658	-1.784	0.074
L2.COP_r	-0.006636	0.018109	-0.366	0.714
L2.CVX_r	-0.029997	0.024368	-1.231	0.218
L2.SPY_r	-0.067779	0.020775	-3.262	0.001
L2.XLE_r	0.014148	0.028677	0.493	0.622
L2.XOM_r	0.015969	0.023415	0.682	0.495

Results for equation XLE_r

	coefficient	std. error	t-stat	prob
const	0.000291	0.000248	1.172	0.241
L1.COP_r	0.012353	0.025701	0.481	0.631
L1.CVX_r	-0.046691	0.034661	-1.347	0.178
L1.SPY_r	0.062370	0.029238	2.133	0.033
L1.XLE_r	-0.025375	0.041008	-0.619	0.536
L1.XOM_r	-0.045301	0.033602	-1.348	0.178
L2.COP_r	-0.030388	0.025721	-1.181	0.237
L2.CVX_r	-0.039309	0.034611	-1.136	0.256
L2.SPY_r	0.014367	0.029508	0.487	0.626
L2.XLE_r	0.021008	0.040731	0.516	0.606
L2.XOM_r	-0.053863	0.033258	-1.620	0.105

Results for equation SPY_r

	coefficient	std. error	t-stat	prob
const	0.000233	0.000175	1.332	0.183
L1.COP_r	0.014031	0.018095	0.775	0.438
L1.CVX_r	-0.026003	0.024404	-1.066	0.287
L1.SPY_r	-0.024006	0.020585	-1.166	0.244
L1.XLE_r	0.002635	0.028872	0.091	0.927
L1.XOM_r	-0.042198	0.023658	-1.784	0.074
L2.COP_r	-0.006636	0.018109	-0.366	0.714
L2.CVX_r	-0.029997	0.024368	-1.231	0.218
L2.SPY_r	-0.067779	0.020775	-3.262	0.001
L2.XLE_r	0.014148	0.028677	0.493	0.622
L2.XOM_r	0.015969	0.023415	0.682	0.495

Results for equation XLE_r

	coefficient	std. error	t-stat	prob
const	0.000291	0.000248	1.172	0.241
L1.COP_r	0.012353	0.025701	0.481	0.631
L1.CVX_r	-0.046691	0.034661	-1.347	0.178
L1.SPY_r	0.062370	0.029238	2.133	0.033
L1.XLE_r	-0.025375	0.041008	-0.619	0.536
L1.XOM_r	-0.045301	0.033602	-1.348	0.178
L2.COP_r	-0.030388	0.025721	-1.181	0.237
L2.CVX_r	-0.039309	0.034611	-1.136	0.256
L2.SPY_r	0.014367	0.029508	0.487	0.626
L2.XLE_r	0.021008	0.040731	0.516	0.606
L2.XOM_r	-0.053863	0.033258	-1.620	0.105

Results for equation XOM_r

	coefficient	std. error	t-stat	prob
const	0.000303	0.000218	1.390	0.165
L1.COP_r	-0.033281	0.022583	-1.474	0.141
L1.CVX_r	-0.012941	0.030456	-0.425	0.671
L1.SPY_r	-0.076478	0.025690	-2.977	0.003
L1.XLE_r	0.037449	0.036033	1.039	0.299
L1.XOM_r	-0.060522	0.029525	-2.050	0.040
L2.COP_r	-0.028990	0.022600	-1.283	0.200
L2.CVX_r	-0.016712	0.030412	-0.550	0.583
L2.SPY_r	0.003271	0.025928	0.126	0.900
L2.XLE_r	0.004973	0.035789	0.139	0.889
L2.XOM_r	-0.062408	0.029223	-2.136	0.033

Correlation matrix of residuals

	COP_r	CVX_r	SPY_r	XLE_r	XOM_r
COP_r	1.000000	0.790282	0.589466	0.855394	0.744536
CVX_r	0.790282	1.000000	0.652165	0.878172	0.829195
SPY_r	0.589466	0.652165	1.000000	0.702569	0.659155
XLE_r	0.855394	0.878172	0.702569	1.000000	0.853645
XOM_r	0.744536	0.829195	0.659155	0.853645	1.000000

A3. Multifactor Risk Exposure Analysis – Comprehensive Results

Out-of-Sample Period

Dep. Variable:	TR	R-squared:	0.075
Model:	OLS	Adj. R-squared:	0.069
Method:	Least Squares	F-statistic:	12.96
Date:	Mon, 24 Jun 2019	Prob (F-statistic):	4.36e-18
Time:	16:15:17	Log-Likelihood:	-2474.5
No. Observations:	1294	AIC:	4967.
Df Residuals:	1285	BIC:	5014.
Df Model:	8		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	0.0070	0.046	0.153	0.878	-0.083	0.097
MKT	0.0709	0.063	1.117	0.264	-0.054	0.195
SMB	0.1136	0.098	1.154	0.249	-0.080	0.307
HML	0.2538	0.128	1.989	0.047	0.004	0.504
RMW	-0.5746	0.160	-3.581	0.000	-0.889	-0.260
CMA	0.6530	0.200	3.269	0.001	0.261	1.045
MOM	-0.4406	0.080	-5.522	0.000	-0.597	-0.284
LOC	-0.1705	0.066	-2.595	0.010	-0.299	-0.042
USD	0.1847	0.106	1.741	0.082	-0.023	0.393

Omnibus:	202.166	Durbin-Watson:	1.945
Prob(Omnibus):	0.000	Jarque-Bera (JB):	2277.122
Skew:	-0.322	Prob(JB):	0.00
Kurtosis:	9.467	Cond. No.	5.15

In-Sample Period

Dep. Variable:	TR	R-squared:	0.219
Model:	OLS	Adj. R-squared:	0.210
Method:	Least Squares	F-statistic:	25.23
Date:	Mon, 24 Jun 2019	Prob (F-statistic):	1.97e-34
Time:	16:14:48	Log-Likelihood:	-1349.3
No. Observations:	729	AIC:	2717.
Df Residuals:	720	BIC:	2758.
Df Model:	8		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	0.0218	0.058	0.377	0.706	-0.092	0.135
MKT	0.0353	0.048	0.736	0.462	-0.059	0.129
SMB	0.0225	0.083	0.271	0.787	-0.141	0.186
HML	-0.0194	0.089	-0.218	0.828	-0.195	0.156
RMW	0.1648	0.168	0.982	0.327	-0.165	0.494
CMA	0.1050	0.182	0.576	0.565	-0.253	0.463
MOM	-0.4421	0.058	-7.598	0.000	-0.556	-0.328
LOC	0.1841	0.052	3.510	0.000	0.081	0.287
USD	-0.2947	0.091	-3.223	0.001	-0.474	-0.115

Omnibus:	61.660	Durbin-Watson:	1.976
Prob(Omnibus):	0.000	Jarque-Bera (JB):	256.019
Skew:	-0.253	Prob(JB):	2.55e-56
Kurtosis:	5.859	Cond. No.	7.81

Total Period

Dep. Variable:	TR	R-squared:	0.173
Model:	OLS	Adj. R-squared:	0.169
Method:	Least Squares	F-statistic:	43.40
Date:	Mon, 24 Jun 2019	Prob (F-statistic):	1.84e-63
Time:	16:12:56	Log-Likelihood:	-2871.3
No. Observations:	1664	AIC:	5761.
Df Residuals:	1655	BIC:	5809.
Df Model:	8		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	0.0426	0.034	1.266	0.206	-0.023	0.109
MKT	0.0317	0.037	0.848	0.397	-0.042	0.105
SMB	0.1407	0.063	2.244	0.025	0.018	0.264
HML	0.1664	0.068	2.453	0.014	0.033	0.299
RMW	-0.1830	0.107	-1.703	0.089	-0.394	0.028
CMA	0.0075	0.122	0.062	0.950	-0.231	0.246
MOM	-0.2748	0.043	-6.332	0.000	-0.360	-0.190
LOC	0.3380	0.039	8.587	0.000	0.261	0.415
USD	-0.0268	0.070	-0.385	0.700	-0.163	0.110

Omnibus:	228.939	Durbin-Watson:	1.969
Prob(Omnibus):	0.000	Jarque-Bera (JB):	2563.851
Skew:	-0.185	Prob(JB):	0.00
Kurtosis:	9.070	Cond. No.	5.99

References

Hendry, D., Juselius, K., 2000. Explaining Cointegration Analysis: Part I and II. Energy Journal.

Diamond, R., 2013. Learning and Trusting Cointegration in Statistical Arbitrage. WILMOTT Magazine, November 2013, Issue 68.

Vidyamurthy, G., 2004. Pairs Trading: Quantitative Methods and Analysis. Wiley.

Gatev, E. G., Goetzmann, W. N. y Rouwenhorst, K. G. 2006. Pairs Trading: Performance of a Relative Value Arbitrage Rule. 2006. Review of Financial Studies 19, 797-827.

Do, B, Faff, R., 2010. Does Simple Pairs Trading Still Work?. Financial Analysts Journal 66, 2010, 83-95.

Mazo, M., Vaquero, M.E., Gimeno, R., 2016. Pairs Trading Strategy and Idiosyncratic Risk. Evidence in Spain and Europe. SSRN Electronic Journal, January 2016.

Rad, H., Kwong Yew Low, R., Faff, R., 2016. The profitability of pairs trading strategies: distance, cointegration and copula methods. 2016. Quantitative Finance 16 (10).

Jacobs, H., Weber, M., 2014. On the Determinants of Pairs Trading Profitability. Journal of Financial Markets.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets 5, 2002, 31–56.

Engelberg, J., Gao, G., and Jagannathan,R. 2009. An Anatomy of Pairs Trading: the Role of Idiosyncratic News, Common Information and Liquidity. Third Singapore International Conference on Finance, 2009.

Lütkepohl, H. 2005. New Introduction to Multiple Time Series Analysis, Chapter 4.3. Wiley, 2005.

Fama, E, French, K. 2014. A Five-Factor Asset Pricing Model. Fama-Miller Working Paper, 2014.