Learning Cointegration for Trading Strategies

Tanya Sandoval

July 25, 2016

Abstract

This report introduces the topic of cointegration and its application to trading strategies. By modelling asset prices with an economic link in common, it is sometimes possible to arrive at a stationary spread whose properties can be used to reduce exposure to systematic risk. The essential elements of cointegration such as stationary and mean-reversion are discussed, as well as some of the statistical tests available to detect this relationship. Simulated and real data examples are provided, the latter focusing on a detected cointegration between Brent Crude and Gasoil futures. Lastly, examples of simple trading strategies applying cointegration are discussed, along with aspects that must be considered when trading under market real conditions.

Contents

1 Introduction

Historically *cointegration* as a concept arose from statistical evidence that many US macroeconomic time series (like GDP, wages, employment, etc.) did not follow conventional econometric theory but rather were described by *unit root processes*, also known as "integrated of order 1" I(1). Before the 1980s many economists used linear regressions on non-stationary time series, which Granger and Newbold showed to be a dangerous approach that could lead to *spurious correlation*. For integrated I(1) processes, Granger and Newbold showed that de-trending does not work to eliminate the problem, and that the superior alternative is to check for cointegration, earning them the Nobel prize.

This report summarises some first learnings of this concept and first-steps at applying it to trading strategies. In particular, the focus is on studying two assets from the commodities space to see if similar properties can be detected as in equities.

- Section 2 provides details for the datasets used throughout the report
- Section 3 introduces stationarity and mean-reversion in time series, which are key elements of cointegration
- Section 4 goes into detail about cointegration and how to test for it, as well as assessing its quality
- Section 5 is then dedicated to its application to trading strategies and assessing their performance in terms of profit and loss (P&L)
- Appendix A summarises some of the mathematical methods involved, such as Multivariate Regression, Autoregressive models (AR(p)), Dickey-Fuller test, optimal lag order and stability conditions
- Appendix B starts to examine cointegration in other energy commodities, such as Italian and Dutch gas

All the relevant scripts to arrive at the results can be found in the project repository "*finalProject/TS*" in the attached USB drive. In particular the ipython notebook *Coint_Brent_Gasoil_v2.ipynb* demonstrates how to run the code, which is omitted in this report for brevity. The project repository is also available online on Github `https://github.com/tsando/CQF/tree/master/finalProject`.

2 Datasets

2.1 Simulated Data

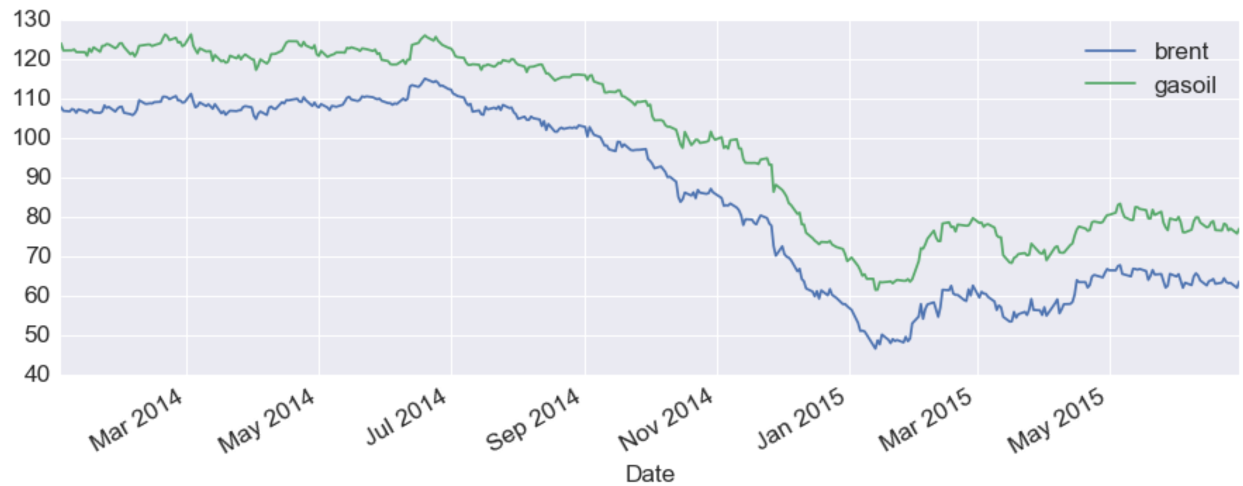
Stochastic processes are used to simulate asset prices. Monte Carlo (MC) techniques are used when relevant and random variables are drawn from the normal distribution $N(\mu = 0, \sigma = 1)$.

2.2 Real Data

- For simplicity, only two financial series are used. As the focus is on commodities, Brent crude and a by-product (Low Sulphur Gasoil) were selected since they were assumed to be good candidates for cointegration given their deep economic link
- The Brent [7] and Gasoil [6] Futures prices traded in the Intercontinental Exchange (ICE) were taken from Quandl's Steven Continuous Series [8] using the *Roll on Last Trading Day with No Price Adjustment* version and the *Settle* field as the closing price on the day

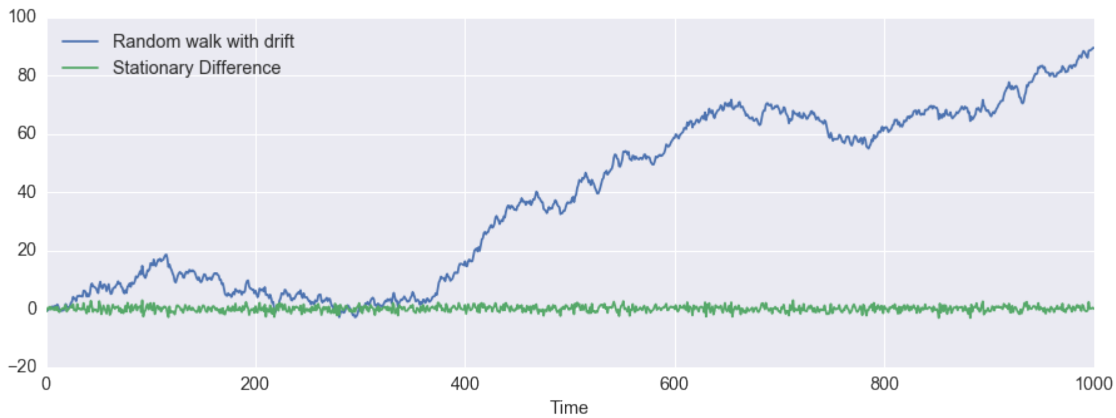
- The two series were then joined to produce a single dataset consisting of daily settlement prices for Brent and Gasoil
- The dataset spans 1.5 'trading years' (equivalent to ~ 252 days). The period selected was Jan-2014 to Dec-2014 for the in-sample testing and Jan-2015 to Jun-2015 for the out-of-sample testing. This was because several sources recommend to use one year of historic data to estimate the cointegration parameters and trade the estimates for a 6-month period, given that the parameters might change over time
- Dates with missing values after joining the two series were removed from the dataset
- Since Gasoil is traded in metric tons and Brent in barrels, the gasoil series was divided by 7.45, which is the ICE conversion factor [5]

The figure below shows the resulting dataset (spanning both in-sample and out-of-sample periods), where the two series indeed seem to be closely related, having very similar trends.



3 Stationarity and Mean-Reversion

Before cointegration is introduced, it is important to understand the concept of *stationarity*. A time series is stationary when the parameters of its generating process do not change over time. In particular, its long-run mean and variance stay constant. This property is fundamental when applying linear regression and forecasting models. Often, processes with a drift or trend, like stock prices, are non-stationary but can be transformed to become stationary. For example, by differencing prices we get returns, which are in general stationary. The figure below shows how a simulated random walk with drift $Y_t = \alpha + Y_{t-1} + \epsilon_t$ can be made stationary by differencing it $\Delta Y_t = Y_t - Y_{t-1} = \alpha + \epsilon_t$



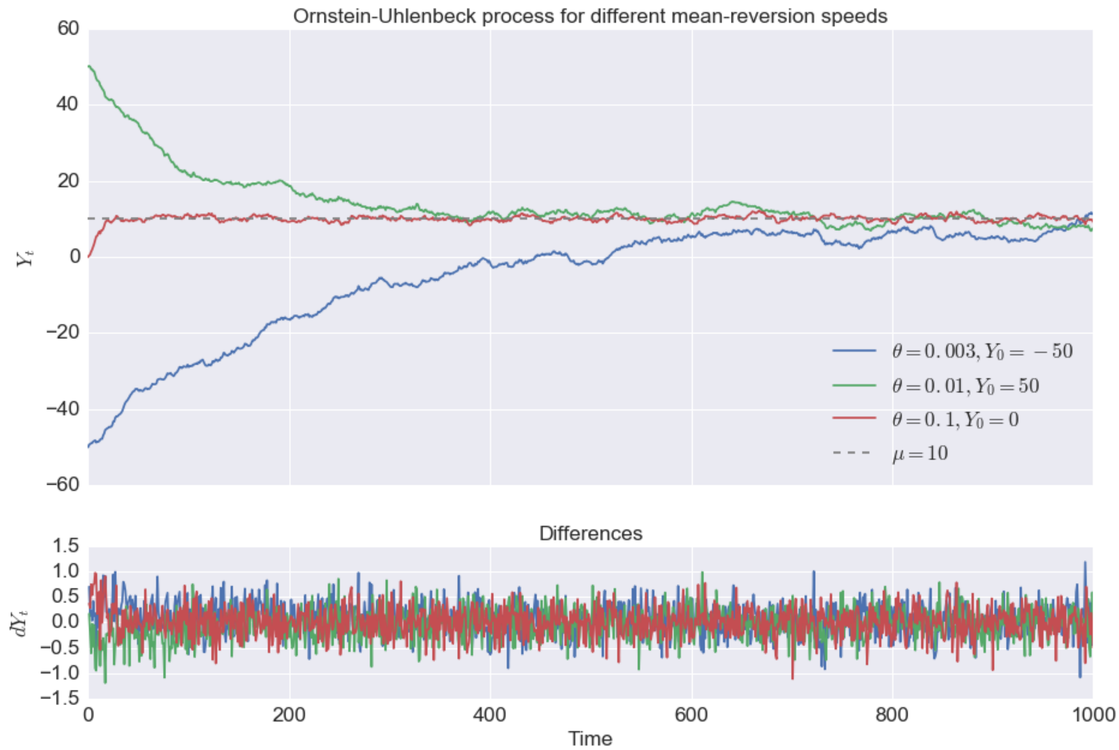
3.1 Mean-Reversion

A stationary series is *mean-reverting* if over time it drifts towards its long-term mean (the historical equilibrium level). A popular model in this category is the Ornstein–Uhlenbeck (OU) process:

$$dY_t = \theta(\mu - Y_t)dt + \sigma dW_t \tag{1}$$

where θ is the speed of reversion, μ is the equilibrium level, σ the variance and W_t is a Wiener Process (Brownian Motion). In a discrete setting this states that the further away the process is from the mean, the greater the 'pull back' to it is. This is in contrast to the random walk above, which has 'no memory' of where it has been at each particular instance of time.

The figure below shows three OU processes with the same mean $\mu = 10$ but different mean-reversion speeds. Indeed it can be noted the one with the highest θ reverts to the mean first. Their differences dY_t are plotted as well and these appear to become stationary significantly faster than the process itself, almost insensitive to the speed θ. Therefore, if we are able to transform a time series to be stationary and mean-reverting, we can design trading strategies using these properties which are more independent of market effects. In a later section we shall see how the OU parameters can be used to design exit/entry thresholds and also assess the *quality* of mean-reversion.



3.2 Tests

We require a more robust method to confirm whether a series is stationary than just by eye. Several statistical tests exist, such as the **Augmented Dickey-Fuller (ADF)** test, Phillips–Perron test, Hurst exponent, Kalman filters, etc. Here we only implement the ADF test and the mathematical details can be found in Appendix A.1.1.

3.2.1 Augmented Dickey-Fuller (ADF)

The ADF test equation implemented was:

$$\Delta Y_t = c_0 + \phi Y_{t-1} + \sum_{k}^{p} \phi_k \Delta Y_{t-k} + \epsilon_t \tag{2}$$

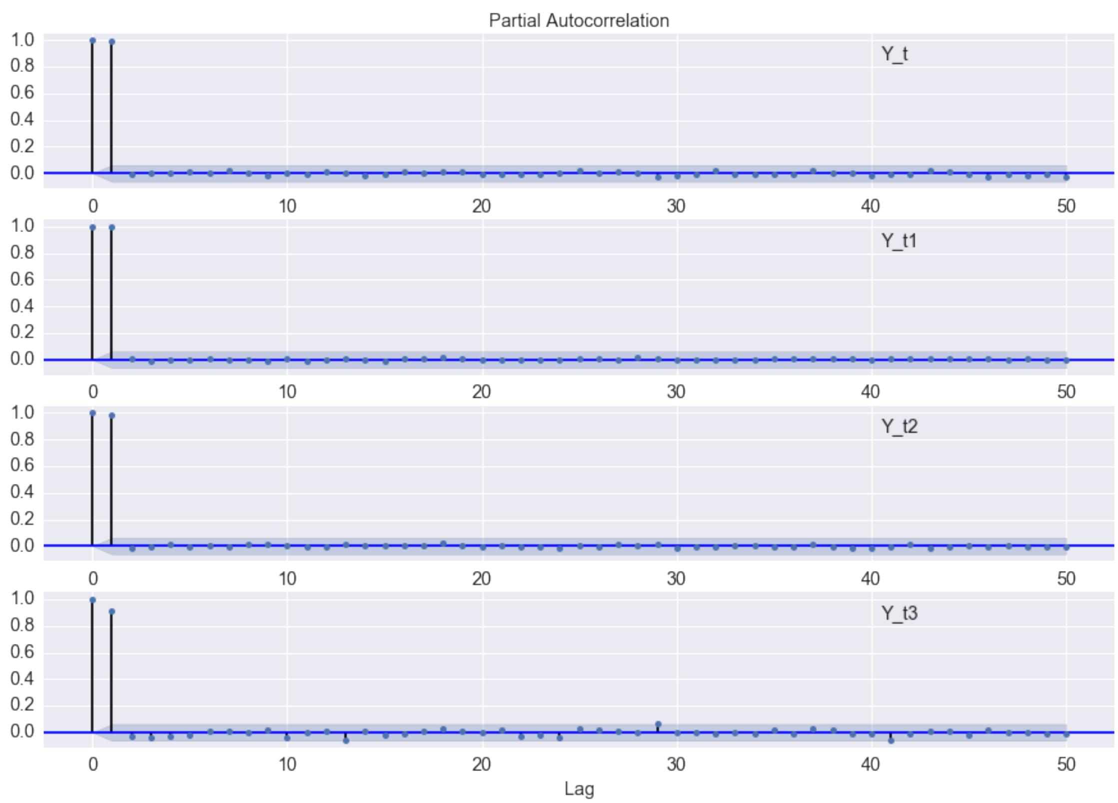
where a time-trend term has not been included due to the nature of financial time series [**?**]. The coefficients are estimated using the familiar linear regression (see Appendix A) whereas the optimal lag order p is discussed below. The python script can be found in *analysis.py*. All the results were validated against the popular python equivalents from the *statsmodels* library [9].

Optimal Lag Selection Choice of lag order can be a difficult problem. Standard approaches use an *information criteria*, such as the Akaike Information Criterion (AIC). However, different methods can lead to different results. Also, keeping more lags can lead to *model overfitting*. In practice, the choice of optimal lag is also evident from the Partial Autocorrelation Function [27] (PACF) since the significant lags would show above confidence limits. Typically for the ADF test it is enough to take p=*1*, however in the interest of exploring this aspect, here we look at the results using AIC and PACF.

- **AIC:** Iterating over different lag orders, the one yielding the lowest AIC value is taken as the optimal lag. For the simulated random-walk-with-drift and OU processes above the results are summarised in the table below.

Process	OU θ	AIC optimal lag
Y_t	0	22
Y_{t1}	0.003	1
Y_{t2}	0.010	10
Y_{t3}	0.100	2

- **PACF:** Given that the optimal lag order from AIC comes out quite high in some cases, we instead use the empirical results from the PACF plot below, where it can be seen only the first lag is well above the 95% confidence band (the first 'spike' represents *p=0*). Given this, we therefore assume it is 'safe' to take *p=1* to carry out the ADF test.



ADF Using as optimal lag *p=1*, we run the ADF test and compare the corresponding t-statistic to the critical values (taken from statsmodels, based on MacKinnon(2010) [12]). The results are summarised in the table below, where we confirm what we expected: the null hypothesis of non-stationary is rejected for all, except for Y_t, which by definition is non-stationary given its drift

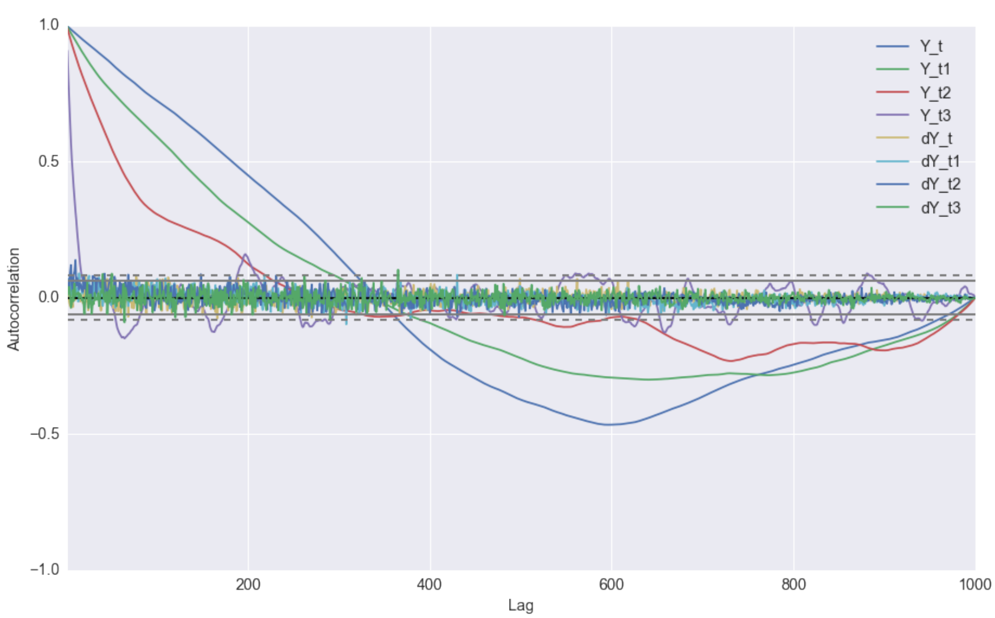
Process	θ	ADF t-stat	5% Crit. Val.	p-value	Stationary	Stable
Y_t	0	0.1093	-2.8644	0.9667	No	Yes
Y_{t1}	0.003	-6.0020	-2.8644	1.65E-07	Yes	Yes
Y_{t2}	0.01	-8.7597	-2.8644	2.69E-14	Yes	Yes
Y_{t3}	0.1	-10.1133	-2.8644	9.87E-18	Yes	Yes
dY_t	0	-31.8476	-2.8644	0.0	Yes	No
dY_{t1}	0.003	-29.4587	-2.8644	0.0	Yes	Yes
dY_{t2}	0.01	-29.3639	-2.8644	0.0	Yes	Yes
dY_{t3}	0.1	-31.3506	-2.8644	0.0	Yes	No

Stability Check To ensure further the reliability of results, a stability check can be done on the estimated coefficients by looking at their eigenvalues within the unit circle (see Appendix A.1.3). The results of the self-implementation are displayed in the table above. All cases were found stable, except dY_t and dY_{t3}. This demonstrates that stationarity does not imply stability. The unstable nature of dY_t may be due to the drift term added $dY_t = \alpha + \epsilon_t$. A close inspection of the problematic root of dY_{t3} shows it is just right on the boundary of the unit circle.

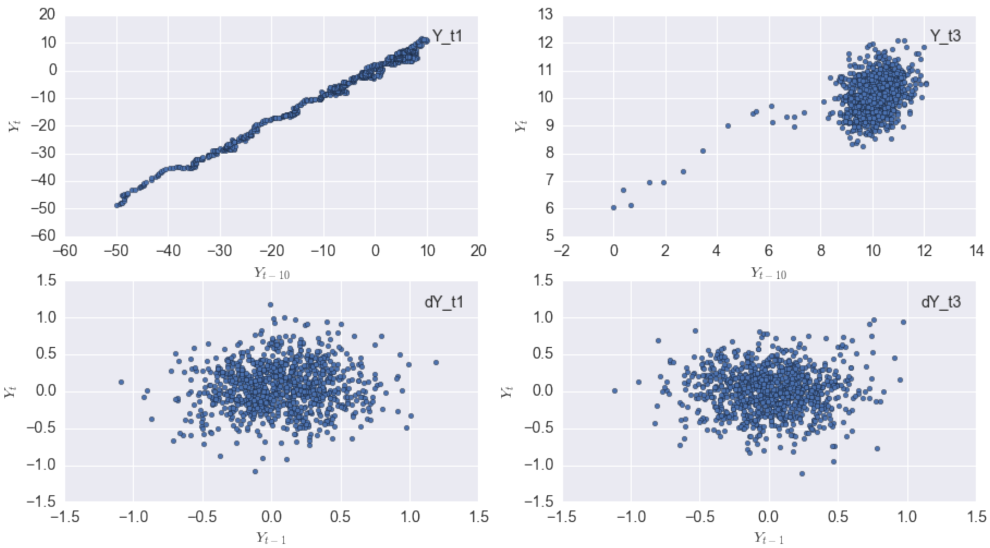
3.3 Other useful plots

In addition to the PACF, when assessing stationarity the below plot types are useful (see references [16] and [17]):

- **Autocorrelation plot:** This shows the autocorrelation function (ACF) at varying time lags. For perfectly stationary series or iid random variables, the autocorrelations should be near zero for all time-lag separations. The horizontal lines displayed in the plot correspond to 95% and 99% confidence levels. Indeed in the example below none of the dY_t processes show significant autocorrelation. Also, the OU process with the highest mean-reversion speed Y_{t3} rapidly loses autocorrelation to its lags, looking stationary to the higher order lags.



- **Lag plot:** this is a scatter plot between the series Y_t and one of its lags Y_{t-p}. Like in the autocorrelation plot, a stationary series would not exhibit any relationship. Below an example is shown for two of the OU processes and their differences. This confirms the results from the autocorrelation plot - the fastest mean-reverting Y_{t3} has little relation to the 10th lag, unlike Y_{t1}, which has a lower speed and a clear relationship. As expected, the stationary differences show little relationship to the lags, even for Y_{t-1}.



4 Cointegration

Having covered key concepts, we now proceed to deep-dive into the topic of cointegration. Two or more time series $\mathbf{Y}_t = (y_{1t}, \ldots, y_{nt})'$ are said to be cointegrated if a linear combination exists which makes the collection 'integrated of order zero' I(0) i.e stationary:

$$\beta'\mathbf{Y}_t = \beta_1 y_{1t} + \cdots + \beta_n y_{nt} \sim I(0) \tag{3}$$

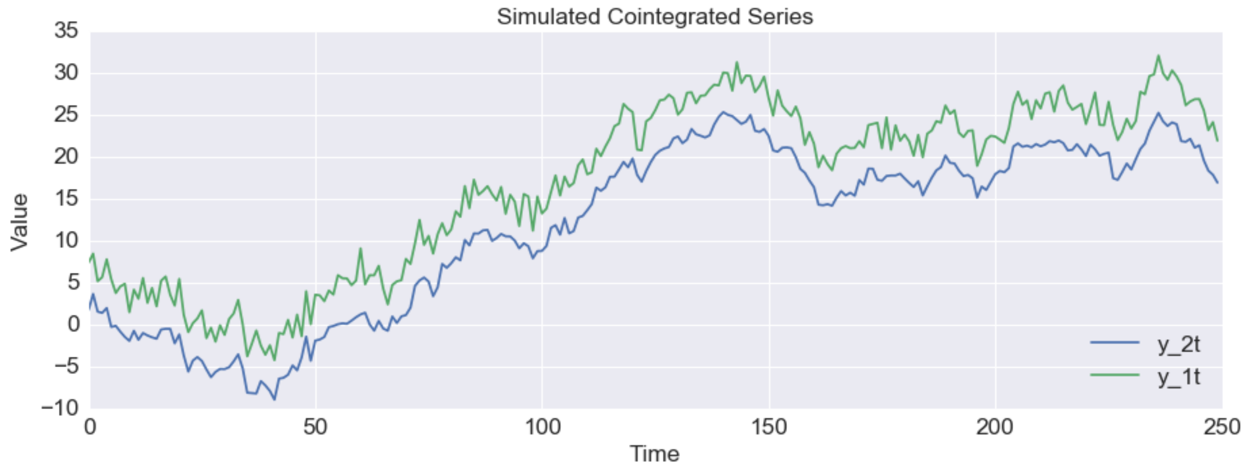
This is known as the *long-run equilibrium model* and is expressed in normalised form as:

$$y_{1t} = \beta_2 y_{2t} + \cdots + \beta_n y_{nt} + e_t \tag{4}$$

where e_t is referred to as the *cointegrating residual* or *spread* and $\beta = (1, -\beta_2, \ldots, -\beta_n)'$ is the *cointegrating vector*. In particular, $e_t \sim I(0)$, i.e. the spread is 'integrated of order 0' - another way to say *stationary*. The concept can be extended to higher orders of integration, although these are more rare.

In real data, cointegration usually exists when there is a deep economic link between the assets and hence these cannot drift too far apart because economic forces will act to restore the long-run equilibrium. The figure below shows a simulated pair of cointegrated assets y_{1t} and y_{2t}, where y_{2t} is defined as an I(1) process. If y_{1t} is supposed to have a strong link to y_{2t}, the price of y_{1t} should vary similarly. This is simulated by shifting up y_{2t} and adding some noise (residual) drawn from a normal distribution, so y_{1t} is defined as the dependent variable and y_{2t} as the independent variable:

$$y_{1t} = y_{2t} + 5.0 + \epsilon_t \tag{5}$$



4.1 Tests

Again, we need a robust method to confirm cointegration. The three main statistical tests are:

- **Engle–Granger two-step method:** This can only be used to test a *single* cointegrating relationship. The steps are:
  - (i) Estimate the cointegrating residual $\hat{e}_t = \hat{\beta}'\mathbf{Y}_t$, e.g. using linear regression
  - (ii) Test $\hat{e}_t$ for stationarity, e.g. using the ADF test, where the hypotheses to be tested are:

$$H_0 : \hat{e}_t = \hat{\beta}'\mathbf{Y}_t \sim I(1) \qquad (\mathrm{nullhypothesis : nocointegration}) \tag{6}$$

$$H_1 : \hat{e}_t = \hat{\beta}'\mathbf{Y}_t \sim I(0) \qquad (\mathrm{alternativehypothesis : cointegration}) \tag{7}$$

- **Johansen test:** Based on maximum likelihood techniques, this allows for more than one cointegrating relationship, but it is subject to asymptotic conditions when the sample size is too small

- **Phillips–Ouliaris test:** Uses a modified version of the Dickey-Fuller distribution to test the cointegrating spread for stationarity. This is a better choice when dealing with small samples

In this project only the Engle-Granger two-step method is considered.

4.1.1 Engle-Granger Two-Step

For our simulated cointegrated series, the cointegration relationship is then represented by the regression:

$$y_{1t} = c + \beta_2' y_{2t} + e_t \tag{8}$$

whose parameters are estimated using OLS. Hence, the estimated cointegrating spread is:

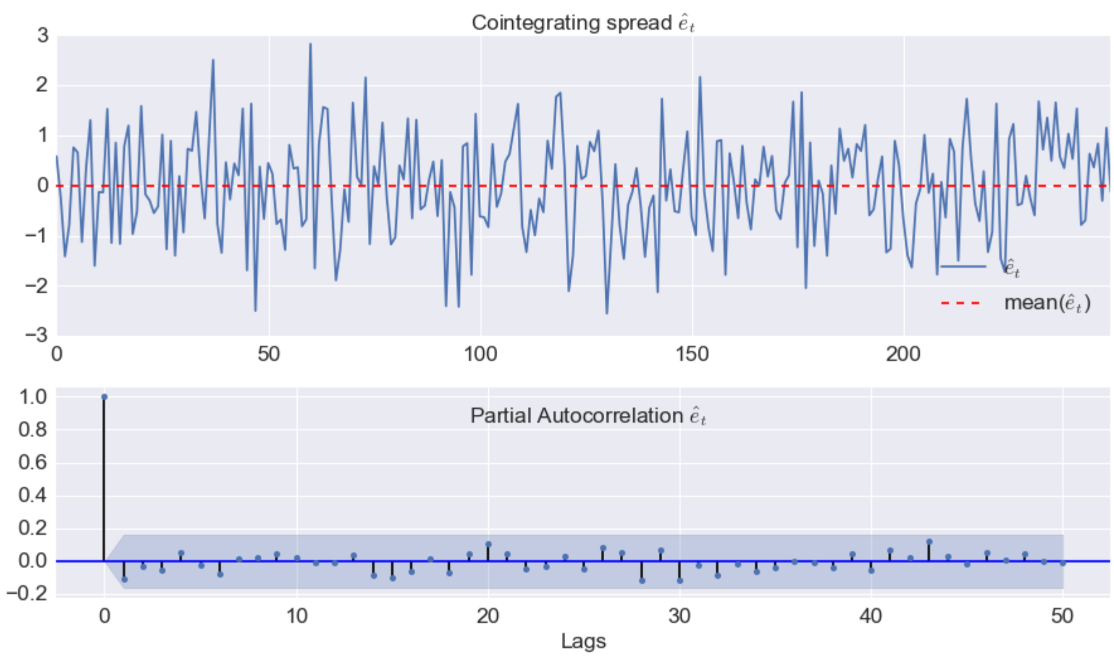
$$\hat{e}_t = y_{1t} - \hat{c} - \hat{\beta}_2' y_{2t} \tag{9}$$

We then test $\hat{e}_t$ for stationarity using ADF. Since the mean of $\hat{e}_t$ is zero, the ADF can be implemented without a constant or trend [1]. Also note that the critical values used are taken as in statsmodels from MacKinnon (2010) [11], whereas other sources suggest to use the Phillips-Ouliaris tabulation. The figure below shows the fitted OLS result:

$$\hat{y}_{1t} = 5.0330 + 1.0052 y_{2t} + \hat{e}_t \tag{10}$$

which is in good agreement with the true parameters we defined.

The estimated $\hat{e}_t$ and its PACF is shown below. The PACF shows the spread is *memoryless*, i.e. no lag order appears significant, as expected from a random process. Conservatively the ADF t-statistic can be computed with one lag. This is -12.051, which is below the 1% critical value (-2.5748), confirming the stationarity of the spread, as expected.



4.2 Error Correction Model (ECM)

Cointegration implies the existence of an Error Correction Model (ECM), which provides an adjustment to the long-run equilibrium from the short-run dynamics. This is particularly useful when modelling non-stationary series, like market prices, which can lead to spurious regression results. Suppose the cointegrated pair is represented by $\mathbf{Y}_t = (y_t, x_t)'$. One can arrive at the ECM result as follows:

- Consider a dynamic regression model to allow for a wide variety of dynamic patterns in the data. This is done by including lags for both x_t and y_t:
$$y_t = \alpha y_{t-1} + \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + \epsilon_t \tag{11}$$
- By knowing the above equation should be consistent with the long-run equilibrium model $y_t = b_0 + b_1 x_t + e_t$, it can be re-written as:
$$\Delta y_t = \beta_1 \Delta x_t - (1-\alpha) e_{t-1} + \epsilon_t \tag{12}$$
where e_{t-1} is the lagged cointegrating spread from the equilibrium model. The parameter $-(1-\alpha)$ can be interpreted as a speed of correction towards the equilibrium level (see next section)
- Since all the variables in the ECM are I(0), OLS can be used to estimate the parameters

4.3 Quality of mean-reversion

If the cointegrating spread e_t is stationary, we could use the OU process to model it:
$$de_t = \theta(\mu_e - e_t)dt + \sigma dW \tag{13}$$
In discrete time this is written as:
$$\Delta e_t = \alpha \mu_e - \alpha e_{t-1} + \epsilon_{t,\tau} \tag{14}$$
where $\alpha = 1 - e^{-\theta\tau}$ and τ is a small period of time. This is in fact the implied ECM representation for e_t. As discussed in Appendix A.1, the above equation can be written in its AR(1) representation as:
$$e_t = \alpha \mu_e + (1-\alpha) e_{t-1} + \epsilon_{t,\tau} \tag{15}$$
Re-writing in terms of $C = \alpha\mu_e$ and $B = 1 - \alpha$ we see it is a simple regression which can be determined with OLS:
$$e_t = C + B e_{t-1} + \epsilon_{t,\tau} \tag{16}$$
In particular, to assess the quality of the spread for trading strategies the parameters of interest are:
$$\theta = -\frac{\ln B}{\tau} \qquad \mu_e = \frac{C}{1-B} \qquad \sigma_{OU} = \sqrt{\frac{2\theta}{1-e^{-2\theta\tau}} Var[\epsilon_{t,\tau}]} \tag{17}$$

- μ_e is the long-run equilibrium level of the OU process
- The mean-reversion speed θ can be translated into a half-life $\tilde{\tau} \propto \ln 2/\theta$, which is the time between equilibrium situations, i.e. when $e_t = \mu_e$. Hence, a high θ (small $\tilde{\tau}$) is desirable to trigger trading signals
- The standard deviation defined as $\sigma_{eq} = \sigma_{OU}/\sqrt{2\theta}$ can be used to plot trading bounds for entry/exit signals as $\mu_e \pm \sigma_{eq}$
- Here we use $\tau = 1/252$ because our data has a daily frequency and the in-sample period covers one trading year, i.e. ~ 252 trading days

4.4 Real Data Example

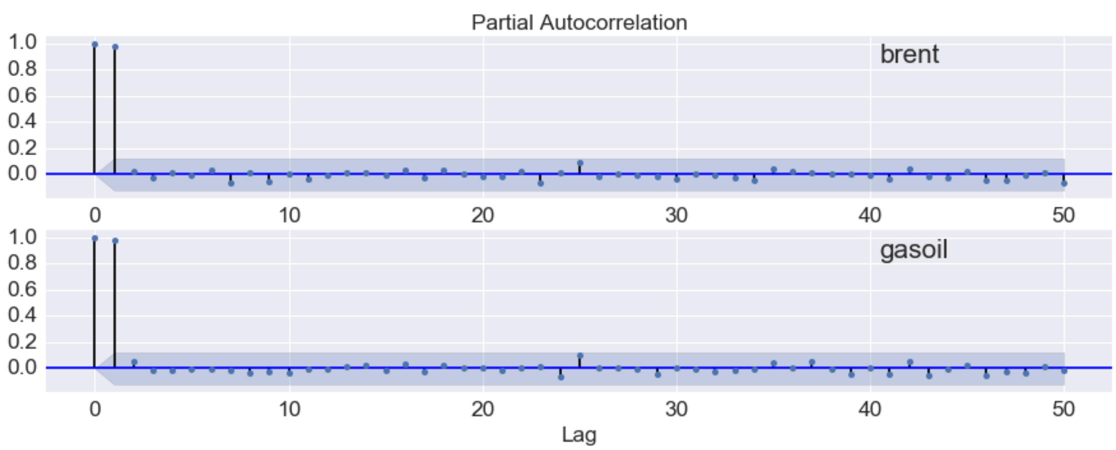
We now move on to apply the concepts to real data, which will highlight some of the challenges that can be encountered.

- We use the **in-sample** dataset for Brent crude and Gasoil (Jan-2014 to Dec-2014) to estimate the cointegration parameters. The **out-of-sample** dataset is later used to backtest trading strategies
- For now we assume Brent is the independent variable 'x_t' and Gasoil the dependent variable 'y_t', and later check if this is accurate via *Granger's causality test*

4.4.1 Stationarity

First we check the individual price series for unit root. We apply the ADF test using one lag only as recommended by their PACF plot below. A drift term (constant) is also included in the test. The results are summarised in the table below. These support the assumption that Brent and Gasoil are I(1) processes whilst their differences are I(0).

Series	ADF t-stat	5% Crit. Val.	p-value	Stationary	Stable
x_t (Brent)	3.5768	-2.8728	1.00	No	No
y_t (Gasoil)	3.9374	-2.8728	1.00	No	No
Δx_t	-18.1437	-2.8728	$< 10^{-30}$	Yes	Yes
Δy_t	-18.9247	-2.8728	0.00	Yes	Yes



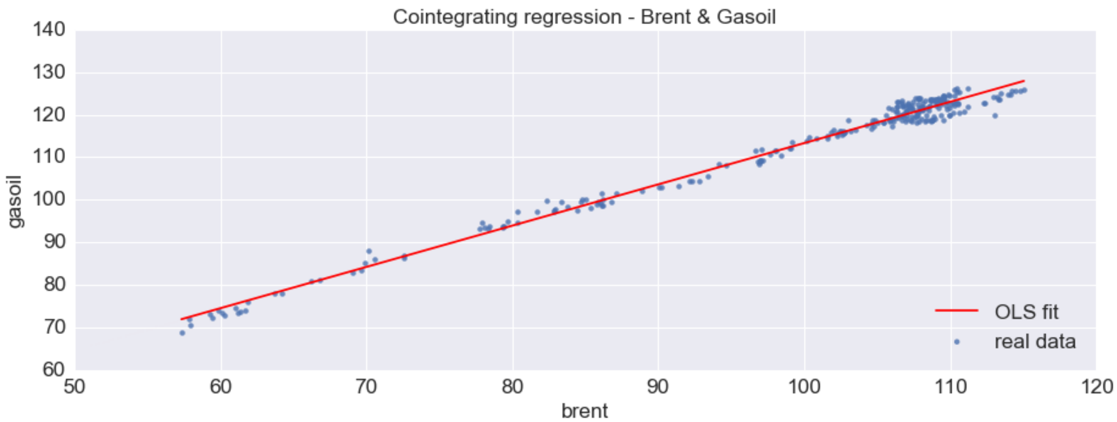
4.4.2 Cointegration

We then proceed to test whether the pair is cointegrated, using the Engle-Granger two-step procedure described in Section 4.1.1.

- The long-run equilibrium model is shown in the figure below. The OLS estimate was:

$$\hat{y}_x = 16.3229 + 0.9699 x_t + \hat{e}_t \tag{18}$$

with goodness-of-fit $R^2 = 0.986$.



- The estimated cointegrating spread $\hat{e}_t$, its distribution and PACF are shown in Figure A. The mean of $\hat{e}_t$ was found to be zero for all practical purposes ($< 10^{-13}$).
- We also want $\hat{e}_t$ to be normally distributed. The histogram shows the spread distribution fitted to a normal distribution with $\mu \sim 0$ and $\sigma = 1.66$. Two tests (*Lilliefors* and *Anderson-Darling*,[13, 18]) were carried out to assess the goodness of fit and neither rejected the null hypothesis that $\hat{e}_t$ is normally distributed
- Next, we note that the PACF plot shows a peculiarity of this spread - it seems to have an AR(3) memory, given the first three lags appear significant. This memory order is actually not desirable to model with the OU process, which is an AR(1) process. It could have an economic explanation, e.g. it could represent cyclicality or storage effects e.g. asset cannot dramatically drop because there are storage/delivery carry costs borne by the sellers. Or it could be from mere fluctuations in the data for this particular dataset, which only spans 1 trading year.
- Next, $\hat{e}_t$ was tested for stationarity using the CADF test. Since the mean is ~ 0 the test was implemented without a constant or trend. Also, as suggested by the PACF plot, the test was ran with 3 lags. The results below show stationary could only be confirmed at the 95% C.L. and not at the 99% C.L.

Series	**ADF t-stat**	**5% Crit. Val.**	**1% Crit. Val.**	**p-value**	**Stationary (95% C.L.)**	**Stable**
e_t	-2.2335	-1.9421	-2.5746	0.0245	Yes	Yes

4.4.3 ECM

The corresponding ECM adjustment was determined to be:

$$\begin{aligned} \Delta y_t &= -0.0861 + 0.6455\Delta x_t - 0.1623 e_{t-1} + \epsilon_t \qquad (\text{with constant fit, } \mathrm{R}^2 = 0.407) \\ \Delta y_t &= 0.6596\Delta x_t - 0.1633 e_{t-1} + \epsilon_t \qquad (\text{no constant fit, } \mathrm{R}^2 = 0.422) \end{aligned} \tag{19}$$

This represents the second-order adjustment to the price, which as we see can be significant.

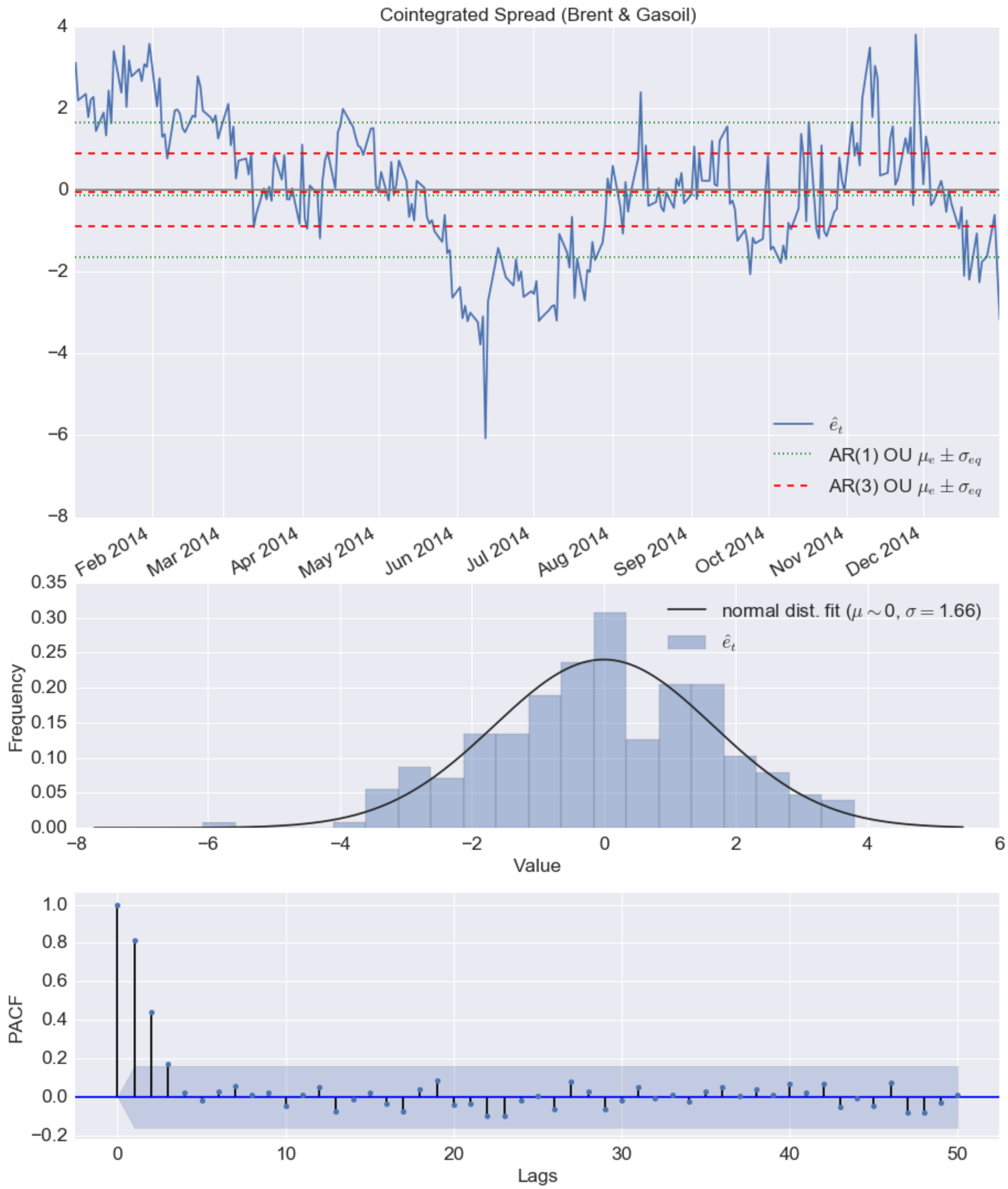


Figure A

4.4.4 Quality of mean-reversion

Given the peculiar PACF of $\hat{e}_t$, which suggests it is an AR(3) process, fitting to the OU process, which is an AR(1) process, might not be suitable. We proceed nevertheless to compare the results between both fits. For AR(3) we take the OU parameters as for AR(1): the constant term equals $\alpha\mu_e$ and the coefficient of the e_{t-1} term is $\alpha = 1 - e^{-\theta\tau}$. The table below shows the results obtained, from which the following trading bounds can be defined:

$$\mu_e \pm \sigma_{eq} = \begin{cases} -0.1255 \pm 1.6541 & \text{AR(1)} \\ -0.0529 \pm 0.8820 & \text{AR(3)} \end{cases} \tag{20}$$

The AR(3) fit yields a lower θ standard error than AR(1), although not by much[1] and yet the θ values are quite different. This perhaps indicates the OU fit is not suitable regardless of the lag order.

Process	θ	$\tilde{\tau}$	μ_e	σ_{OU}	σ_{eq}	$s.e._{\theta}$
AR(1)	49.1253	0.0141	-0.1255	16.3957	1.6541	7.3850
AR(3)	258.9327	0.0027	-0.0529	20.0723	0.8820	5.6638

4.5 Granger Causality

When implementing the cointegrating regression to estimate e_t, an assumption must be made about which the dependent variable is. It is actually important to determine the best choice for this as it can largely influence the results.

Grange causality means that lags of x_2 have a statistically significant effect on the current value of x_1, taking lags of x_1 into account as regressors. We reject the null hypothesis H_0 that x_2 does not Granger cause x_1 if the pvalues are below a desired size of the test. The ready implementation from statsmodels [15] was used, which uses the F distribution critical values. Running the test over the in-sample Brent-Gasoil data using up to 3 lags (which should be enough given theirs PACF plot) yields the results in the table below. This confirms that taking Brent as the independent variable is acceptable. The hypothesis of Gasoil causing Brent moves is rejected too, which further supports the result.

	F test p-value		
H_0	1 lag	2 lags	3 lags
Brent doesn't cause Gasoil	0.0005	0.0001	<10^-4
Gasoil doesn't cause Brent	0.2315	0.2986	0.4797

[1] To calculate the standard error, the simplified formula for error propagation in [22] was used

5 Trading Strategies

Below we describe a few simple strategies which exploit the cointegration properties so far discussed. The first step is to construct a *stationary portfolio* of separately interrelated assets. For the demonstrations below however, we assume this portfolio is limited to the 2 assets from before: Brent and Gasoil.

5.1 Regime changes

Tests for cointegration assume that the cointegrating vector is constant during the period of study. In reality, it is common for the relationship of the underlying variables to change, e.g from changes in their *fundamental factors*. This is especially likely to be the case if the sample period is long. Hence, it should not be assumed that because the assets have passed a cointegration test historically, they will continue to remain cointegrated. Practitioners' advice is to estimate using one year of historic data and trade the estimates for a 6-month period. This is then a good introduction to *backtesting*.

5.2 Backtesting

Backtesting is the process of testing a trading strategy or model on historical data to gauge its effectiveness. When we 'backtest a model', the results achieved are highly dependent on the tested period and this can cause the strategy to fail in the future due regime changes or *model overfitting*. To alleviate this, the test data is usually split into *in-sample* and *out-of-sample* (a 'rule-of-thumb' is to use an 80-20 split). The model/strategy is then fitted/tested to the in-sample data *only* and then tested on the out-of-sample data, which was 'unseen'. This process provides a better way to assess the true performance of a strategy.

5.3 Naive Beta-Hedging Strategy

Factor models are a way of explaining the returns of one asset via a linear combination of the returns of other assets. This can be expressed as a simple linear regression:

$$\Delta Y = \alpha + \beta_1 \Delta X_1 + \beta_2 \Delta X_2 + \cdots + \beta_n \Delta X_n + \epsilon_t \tag{21}$$

We can interpret the betas as the *exposure* of asset Y to the other assets and α as the market-neutral excess return:

$$\alpha = \mathrm{E}[\Delta Y - \beta_1 \Delta X_1 + \beta_2 \Delta X_2 + \cdots + \beta_n \Delta X_n] \tag{22}$$

since $\mathrm{E}[\epsilon_t] = 0$. For the case of Brent and Gasoil only this would be:

$$\Delta Y_{gasoil} = \alpha + \beta \Delta X_{brent} + \epsilon_t \tag{23}$$

i.e. we could take a short position in Brent equal to $\beta \Delta X_{brent}$ to try to eliminate the *risk* in our Gasoil position. On average, we would expect our returns to be:

$$\mathrm{E}[\Delta Y_{gasoil} - \beta X_{brent}] = \alpha \tag{24}$$

For the **in-sample** data, running the above regression would actually yield negative returns, although less severe than just going long on Gasoil:

$$\begin{aligned} \mathrm{E}[\Delta Y^{is}_{gasoil} - 0.6282 \Delta X^{is}_{brent}] = &\quad \alpha = -0.0915 \qquad (R^2 = 0.354) \\ \mathrm{E}[\Delta Y^{is}_{gasoil}] = &\quad \alpha = -0.2148 \qquad \text{(no fit)} \end{aligned} \tag{25}$$

This shows the estimated beta is not constant as we walk forward in time. As such, the short position we took out in Brent is not perfectly hedging our portfolio. Another is that additional factors or assets should be included into the model. Surprisingly, our performance in the **out-of-sample** seems to be better but that was just mere 'luck' from positive fluctuations in the data, and still worse off than just going long on Gasoil over this period:

$$\begin{aligned} \mathrm{E}[\Delta Y^{os}_{gasoil} - 0.6282\Delta X^{os}_{brent}] &= \mathrm{E}[\alpha_t] = 0.02231 \qquad \text{(in-sample fit)} \\ \mathrm{E}[\Delta Y^{os}_{gasoil}] &= \mathrm{E}[\alpha_t] = 0.05806 \qquad \text{(no fit)} \end{aligned} \tag{26}$$

5.4 Pairs Trading Strategy

Moving to a more sophisticated approach with cointegration, we now look into 'pairs trading'. For a pair of assets, if they are cointegrated, e.g. because they belong to the same 'sector', they are likely to be exposed to similar market factors. Occasionally their relative prices will diverge due to certain events, but eventually will revert to the long-run equilibrium. Hence, positions are taken relative to where the **cointegrating spread** $\hat{e}_t$ is with respect to this equilibrium. The objective is to hedge the position from price level dynamics (systematic risk). The P&L will then be driven by the quality of mean-reversion (Section 4.4.4). For Gasoil and Brent, we estimated the **in-sample** relationship as:

$$Y^{is}_{gasoil} = 0.9699X^{is}_{brent} + 16.3229 + \hat{e}^{is}_t \tag{27}$$

Under the assumption that the relationship holds, we use the same beta and constant to estimate the **out-of-sample** spread:

$$\hat{e}^{os}_t = Y^{os}_{gasoil} - 0.9699X^{os}_{brent} - 16.3229 \tag{28}$$

This is then what we use to **backtest** the strategy.

5.4.1 Bounds from OU process

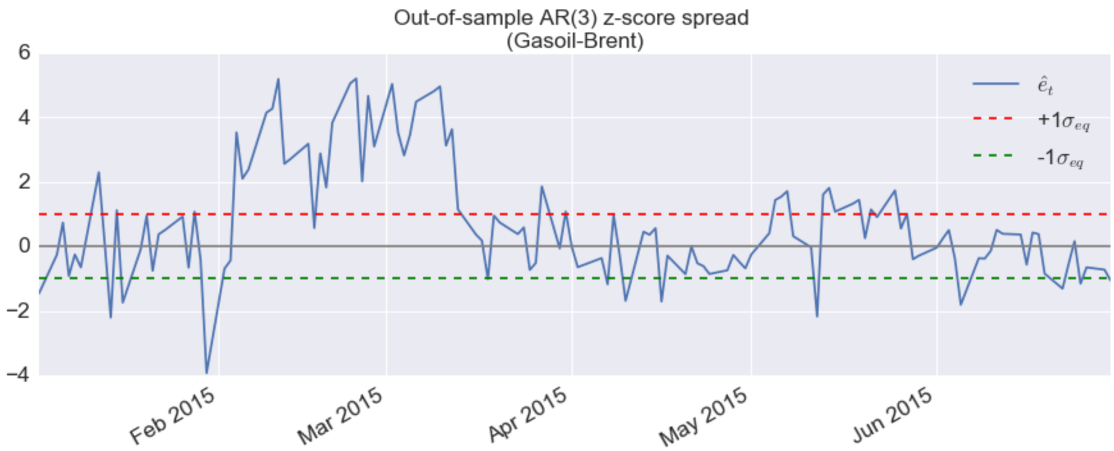
As we saw, we could define thresholds for trading the spread by fitting to the OU process and using the fitted mean μ_e and relative standard deviation σ_{eq}. Assuming normality of the spread $e_{t,\tau\to\infty} \sim N(\mu_e, \sigma^2_{eq})$, a z-score could be used to 'normalise' the spread as:

$$z = \frac{\hat{e}_t - \mu_e}{\sigma_{eq}} \tag{29}$$

With the z-score, a simple strategy would then be:

- Go "Long" the spread when the z-score is below -1.0, as the expectation is that it will rise (this means we are longing Y_{gasoil} and shorting X_{brent})
- Go "Short" the spread when the z-score is above 1.0, as the expectation is that it will fall (this means we are shorting Y_{gasoil} and longing X_{brent})
- Exit positions when the z-score is near zero, e.g. within $[-0.1, 0.1]$

The figure below shows the out-of-sample z-score spread with the AR(3) OU entry/exit bounds marked.



The corresponding P&L is shown in Figure B for both the in-sample and out-of-sample periods. A positive performance is observed throughout, although with prolonged drawdown periods. In addition, note this does not include **transaction costs** and other market effects which could certainly add up to an overall loss. The P&L using the AR(1) bounds is also shown, where it is clear it is worse off than AR(3). Both however are better off than having simply gone long in Gasoil or Brent - this is taken as the market 'benchmark'. Given the positive results, it would be interesting to see the P&L from extending to more than 2 assets cointegrated with oil, with a strategy that trades multiple cointegrating spreads at the same time, or a single spread composed of more factors. In this case, the **Johansen test** would be more suitable, although this is not covered here.

5.4.2 Optimal bounds

Although earlier the Gasoil-Brent spread passed the Anderson-Darling test for normality, we also saw it had memory AR(3). Hence, it is perhaps not optimal to use an OU fit to define the bounds. The P&L plot also suggests the performance is sensitive to where the bounds are. This motivates the search for the *optimal bounds*, which maximise the P&L or any other metric of interest, e.g. Sharpe ratio or number of trades. Optimisation techniques can be used to test different combinations of $\mu_e \pm \sigma_{eq}$ in the in-sample data. A dynamic/rolling bound could even be defined and an optimisation could be ran to find the optimal period for the rolling window, which could be different between μ_e and σ_{eq} given their different time-scales. However, the impact on performance in the out-of-sample would have to be addressed given that this introduces additional parameters into the model, potentially causing model-overfitting.

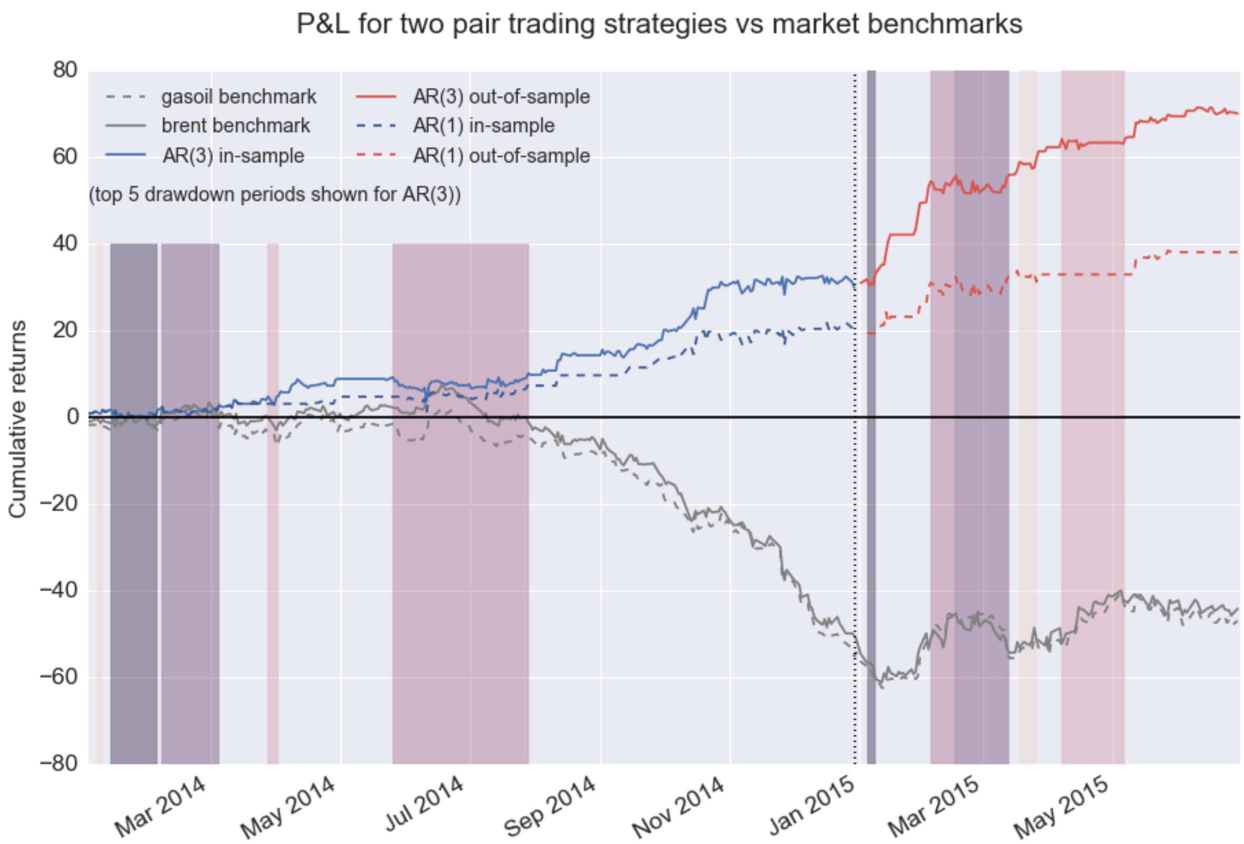
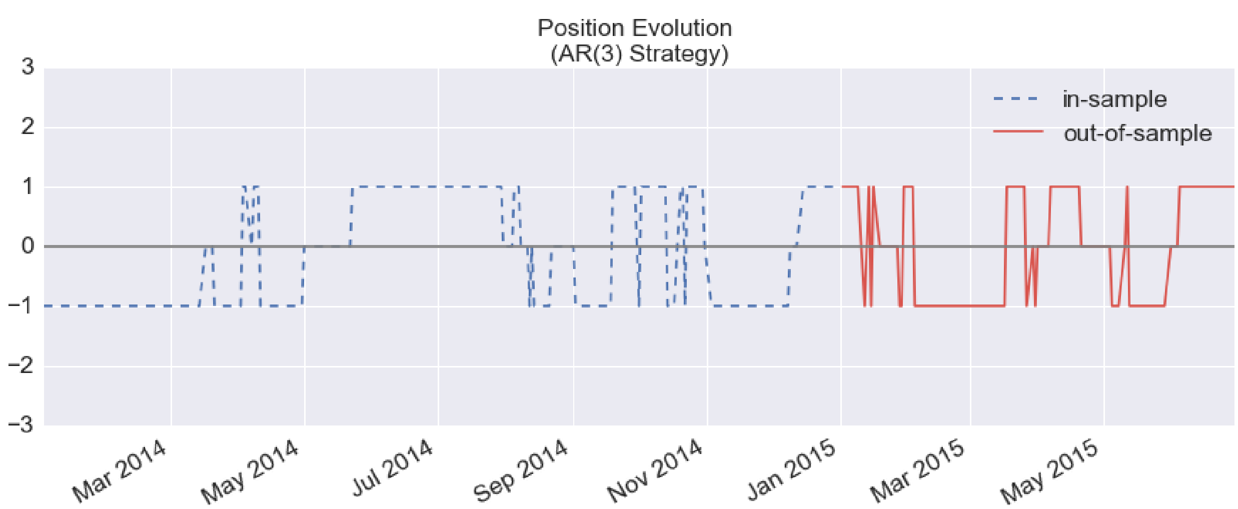

P&L for two pair trading strategies vs market benchmarks
gasoil benchmark
brent benchmark
AR(3) in-sample
AR(3) out-of-sample
AR(1) in-sample
AR(1) out-of-sample
(top 5 drawdown periods shown for AR(3))
Cumulative returns
Mar 2014
May 2014
Jul 2014
Sep 2014
Nov 2014
Jan 2015
Mar 2015
May 2015

Figure B


Position Evolution
(AR(3) Strategy)
in-sample
out-of-sample
Mar 2014
May 2014
Jul 2014
Sep 2014
Nov 2014
Jan 2015
Mar 2015
May 2015

6 Conclusion

Through this project I learned for the first time how to apply cointegration to time series data and design trading strategies around it that could potential generate business value. The P&L obtained from trading a simple Brent-Gasoil spread in the out-of-sample was surprisingly good, given the low performance one often finds at this step already. However, I feel the validity of results for real trading is debatable due to the following:

- Transaction costs and market effects weren't really considered and this could have a large effect given the high number of positions taken (see position plot above)
- The AR(3) memory in the Brent-Gasoil spread wasn't well understood and it wasn't clear whether this was a problem with the data or truly economic effects
- A thorough data quality assessment on the Quandl dataset wasn't done and this had some adjustments applied (e.g. roll corrections) which would need to be understood
- The strategy wasn't backtested over other out-of-sample periods to see if the 'good performance' holds

In addition, given the time limitations it wasn't possible to deep-dive into some interesting aspects which I would have liked to understand better, such as:

- Designing trading strategies for more than 2 assets in the commodities space. In particular it would be good to gain some experience with the Johansen procedure (currently not implemented in python's statsmodels library)
- Optimisation of P&L metrics, like Sharpe ratio and trading bounds
- Getting additional plots and metric to assess the P&L, such as those found in the Pyfolio package [20]

Regardless, cointegration should not be assumed to yield always meaningful results, specially when working with real data. Some frequent issues to be aware of are:

- Statistical tests may find cointegration when there is actually *none*. Moreover, the tests are quite sensitive to how the regression is formulated, e.g. including a deterministic trend or drift term could yield a different result, as well as different number of lags. Conversely, tests can also fail at detecting it
- Cointegration can be a temporary effect, and testing needs to be constantly done to confirm the stationarity of the spread, yet, which period to use for testing is somehow arbitrary
- The 'long-term equilibrium' is a relative concept and there is no 'best' way to define it
- Using cointegration for forecasting can lead to large errors, specially when forecasting series which aren't stationary

References

[1] *Modeling Financial Time Series with S-PLUS - Chapter 12 - Cointegration* by Zivot, E. and Wang J. (2006)

[2] CQF lectures on *Cointegration and Statistical Arbitrage*, Diamond, R. (2016)

[3] *Learning and Trusting Cointegration in Statistical Arbitrage*, Diamond, R., Wilmott Magazine (2013)

[4] *Personal Github repository* https://github.com/tsando/CQF/tree/master/finalProject

[5] https://www.theice.com/publicdocs/futures/ICE_Gas_Oil_Crack.pdf

[6] https://www.theice.com/products/34361119/Low-Sulphur-Gasoil-Futures

[7] https://www.theice.com/products/219/Brent-Crude-Futures

[8] https://www.quandl.com/data/SCF/documentation/about

[9] http://statsmodels.sourceforge.net/

[10] http://statsmodels.sourceforge.net/stable/vector_ar.html#lag-order-selection

[11] http://statsmodels.sourceforge.net/stable/generated/statsmodels.tsa.stattools.adfuller.html#statsmodels.tsa

[12] http://statsmodels.sourceforge.net/stable/generated/statsmodels.tsa.stattools.adfuller.html#statsmodels.tsa

[13] http://statsmodels.sourceforge.net/notebooks/generated/statsmodels.stats.diagnostic.normal_ad.html

[14] http://matthieustigler.github.io/Lectures/Lect2ARMA.pdf

[15] http://statsmodels.sourceforge.net/stable/generated/statsmodels.tsa.stattools.grangercausalitytests.html#sta

[16] http://pandas.pydata.org/pandas-docs/stable/visualization.html#autocorrelation-plot

[17] https://github.com/pydata/pandas/blob/master/pandas/tools/plotting.py

[18] http://statsmodels.sourceforge.net/devel/generated/statsmodels.stats.diagnostic.kstest_normal.html

[19] Quantopian Lectures - quantopian.com/lectures

[20] https://github.com/quantopian/pyfolio/blob/master/pyfolio

[21] https://en.wikipedia.org/wiki/Information_criterion

[22] https://en.wikipedia.org/wiki/Propagation_of_uncertainty#Simplification

[23] https://en.wikipedia.org/wiki/Autoregressive_model

[24] https://en.wikipedia.org/wiki/Vector_autoregression

[25] https://en.wikipedia.org/wiki/Augmented_Dickey%E2%80%93Fuller_test

[26] https://en.wikipedia.org/wiki/Dickey%E2%80%93Fuller_test

[27] http://nl.mathworks.com/help/econ/autocorrelation-and-partial-autocorrelation.html

Appendix

A Multivariate Regression

Also known as 'generalised linear model', it generalises linear regression to multiple input variables (regressors) and n observations. It is best expressed in matrix form as:

$$Y = X\beta + \epsilon \tag{30}$$

where Y is a vector representing the endogenous (dependent) variables, X is a matrix representing the exogenous (independent) variables, β is the coefficients vector and ϵ the residuals vector. The OLS method, which minimises the sum of squared residuals via the Maximum Likelihood Estimation method (MLE), is used to estimate the parameters:

$$\hat{\beta} = (X'X)^{-1}X'Y \tag{31}$$

$$\hat{\epsilon} = Y - X\hat{\beta} \tag{32}$$

The covariance matrix of the residuals estimate is:

$$\hat{\Sigma} = scale \times \sum \epsilon\epsilon' \tag{33}$$

where the scale equals $1/n$ using the MLE estimator or $1/(n-kp)$ using the OLS estimator for a model with k variables and p lags.

The covariance matrix for the coefficients is:

$$(XX')^{-1} \otimes \hat{\Sigma} \tag{34}$$

where $\otimes$ is the *Kronecker product*.

The Log Likelihood Function for OLS is:

$$\log(L) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log|\hat{\Sigma}| - \frac{n}{2} \tag{35}$$

The variance of the residuals and parameters are therefore the diagonal elements of the corresponding covariance matrices, from which the standard errors can be calculated. These are the conventions used in the script *analysis.py*. Additional mathematical details can be found in reference [24].

There are several assumptions about the nature of the variables in multivariate regression. In particular, these should be *stationary* and the residual ϵ homoscedastic (with finite variance) and normally distributed. The main applications of the multivariate regression are:

- Autoregression Models - these can be used to forecast and test or model stationary time series
- Error Correction Model - these are used to model series which aren't stationary or that have stochastic trends, like prices

A.1 Autoregression Models - AR(p)

Also referred to as AR(p) where p is the lag order, it is simply a linear regression on a time series and its lagged (past) values:

$$Y_t = c + \sum_{i=1}^{p} \phi_i Y_{t-i} + \epsilon_t \tag{36}$$

where c is a constant (also known as the drift term), ϕ_i are the parameters of the model and ϵ_t is the error term. Whether to exclude the constant c or not depends on the nature of what we are trying to model. Computationally, the model can be fitted in one go by using the OLS method described above with a special matrix formulation. Example code for this can be found in *analysis.py* in the project repository.

An AR(p) system can be re-written in terms of differences, which is how it is commonly expressed in some cases, for example in the ADF test. See more details in references [24] and [23].

A.1.1 Dickey-Fuller Test and ADF

The *Dickey-Fuller test* examines the null hypothesis of whether a unit root is present in the autoregressive model AR(p) of a time series. For example, a simple AR(1) model is:

$$Y_t = \beta Y_{t-1} + \epsilon_t \tag{37}$$

If $\beta = 1$ the series is said to have a 'unit root' and hence is non-stationary. The equation can be re-written as:

$$\Delta Y_t = (\beta - 1) Y_{t-1} + \epsilon_t = \phi Y_{t-1} + \epsilon_t \tag{38}$$

where $\phi = \beta - 1$. Hence, testing for unit root is equivalent to testing $\phi = 0$. The value of the test statistic $\frac{\hat{\phi}}{\text{std.err}(\hat{\phi})}$ is then compared to the relevant critical values for the Dickey-Fuller distribution. If found lower, then the null hypothesis $\phi = 0$ is rejected and the series can be considered stationary. There are three main versions of the test depending on whether drift and/or time-dependent terms are included:

- Test for a unit root: $\Delta Y_t = \phi Y_{t-1} + \epsilon_t$
- Test for a unit root with drift: $\Delta Y_t = c_0 + \phi Y_{t-1} + \epsilon_t$
- Test for a unit root with drift and deterministic time-trend: $\Delta Y_t = c_0 + c_1 t + \phi Y_{t-1} + \epsilon_t$

Each version of the test has its own critical values which depend on the size of the sample.

Which version to use is not straightforward and the wrong choice can lead to wrong result. In general, financial time series exclude the time-trend. There is an extension to the test referred to as the *Augmented Dickey–Fuller test* (ADF), which removes autocorrelation effects by including lagged difference terms $\phi_p \Delta Y_{t-p}$. The optimal lag order could then be determined from an information criteria (see below).

As this belong to the family of generalised linear models, the parameters can be estimated using the multivariate regression described above. Additional details on this topic can be found in references [25], [26].

A.1.2 Optimal Lag Order

To select the optimal lag order, one approach uses the Akaike Information Criterion (AIC). Iterating over different lag orders, the one yielding the lowest value of AIC is selected. Statsmodels suggests to try up to a maximum lag order of $12 * (n/100)^{1/4}$ where n is the number of observations. There are different definitions of AIC used - we use the same as in statstools [10], which has different definitions for AR(p) and the ADF test as:

$$AIC = \log|\hat{\Sigma}| + 2\frac{1+k}{n} \qquad \text{(AR(p) model)} \tag{39}$$

$$AIC = -2\log(L) + 2k \qquad \text{(ADF test)} \tag{40}$$

where k is the number of estimated parameters. Other information criteria can be used, see for example reference [21].

A.1.3 Stability Condition

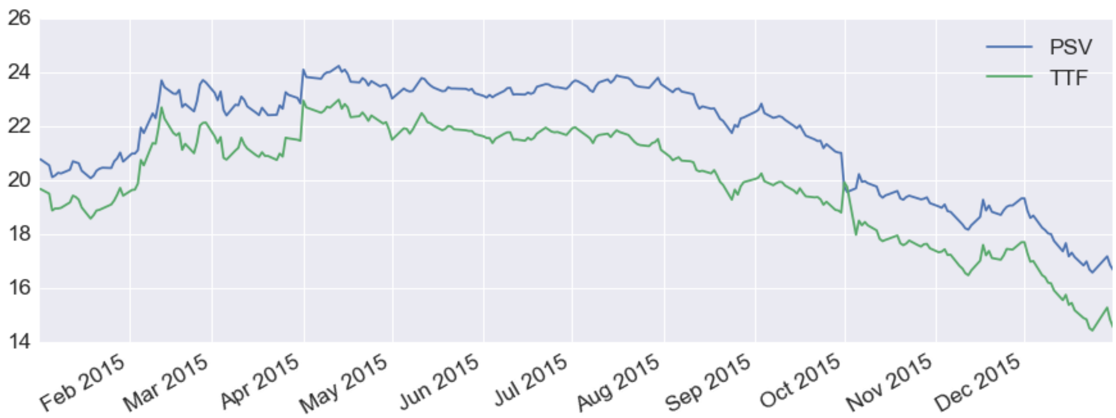
For an AR(p) system, it is required for the eigenvalues of the estimated coefficients to be inside the unit circle (<1):

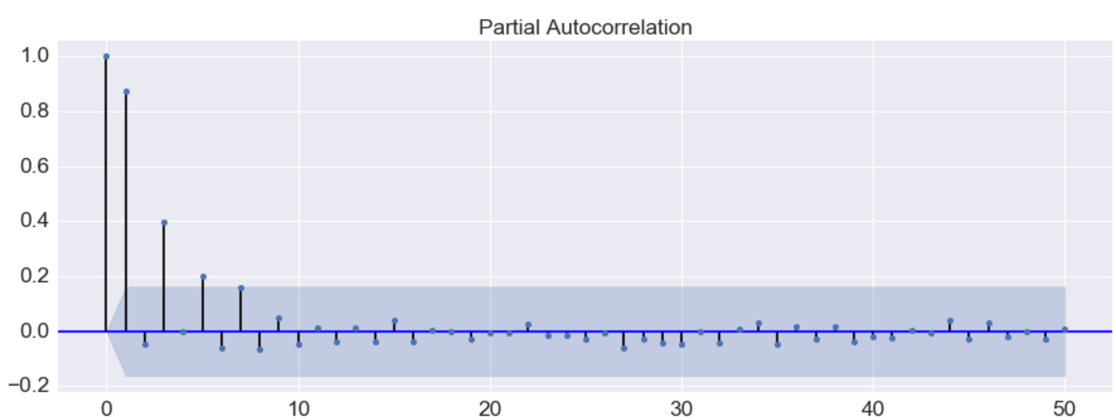
$$|\lambda I - \hat{\beta}| = 0 \tag{41}$$

This is equivalent to requiring the roots of the characteristic polynomial of the AR(p) system to be outside the unit circle - see [21].

B Cointegration between Italian and Dutch Gas

A similar study was started using the price series of Italian (PSV) and Dutch Gas (TTF) futures, however, due to time limitations this wasn't finished. In any case one could already see from the spread plot that the quality wasn't as good as that of Brent-Gasoil (subject to data quality assessment). Some of the related plots are shown below. Note that the spread memory noticed in Brent-Gasoil seems more severe in this case.





Credit Valuation Adjustment for an Interest Rate Swap

Tanya Sandoval

July 24, 2016

Abstract

In this report we demonstrate some of the techniques currently used to calculate Credit Valuation Adjustments for fixed income instruments. We take the hypothetical scenario of an Interest Rate Swap entered by two counterparties. Through the simulation of forward rates using the HJM model on recent data and default probabilities using Credit Default Swaps spreads, we arrive at the fair price of the risk taken by counterparty A in entering the swap with counterparty B.

Contents

1 Introduction

We wish to model the scenario of an Interest Rate Swap (IRS) entered by two counterparties 'A' and 'B'. The IRS is assumed to be written on a 6M LIBOR L_{6M} variable rate, with notional $N = 1$. To calculate the Credit Valuation Adjustment the main inputs would be:

- Default Probabilities (PDs)
- Forward LIBORs
- Discount Factors (DFs)

Below we discuss each of these aspects separately. All the relevant scripts and spreadsheets to arrive at the results can be found in the project repository "*finalProject/CVA*" in the attached USB drive. In particular the ipython notebook *CVA.ipynb* demonstrates how to run the code, which is omitted in this report for brevity. The project repository is also available online on Github `https://github.com/tsando/CQF/tree/master/finalProject`.

2 Default Probabilities

The Default Probabilities can be implied from Credit Default Swaps (CDS) spreads using the bootstrapping technique.

2.1 CDS bootstrapping

The bootstrapping formula used was:

$$P(T_N) = \frac{\sum_{n=1}^{N-1} D(0,T_n)[LP(T_n - 1) - (L + \Delta t_n S_N)P(T_n)]}{D(0,T_N)(L + \Delta t_n S_N)} + \frac{P(T_{N-1})L}{(L + \Delta t_N S_N)} \tag{1}$$

where P is the survival probability, $L = 1 - R$ is the expected loss calculated from the recovery rate R, Δt is the payment frequency, S is the CDS spread and D is the discount factor

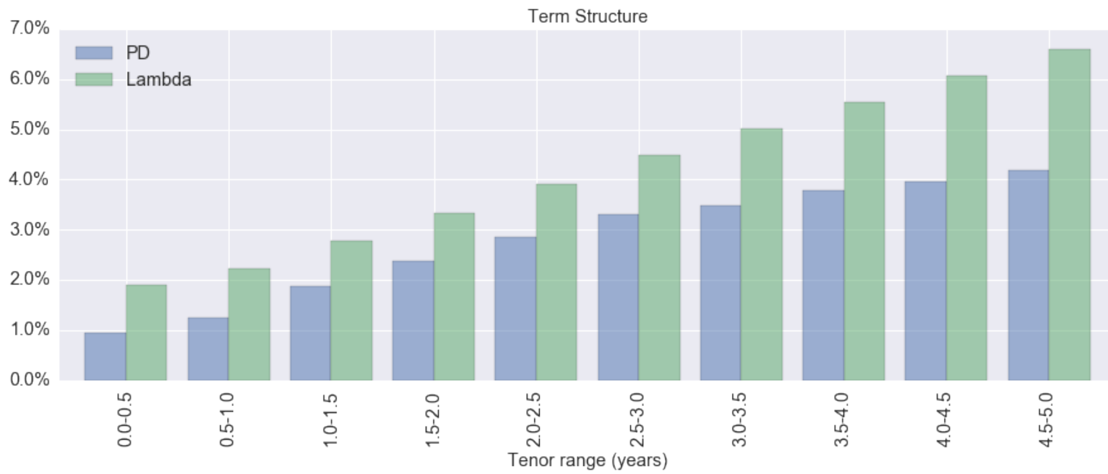
- The counterparty B in this case was chosen to be an airline company - AirFrance ("AIRF")
- The CDS spreads were taken from Reuters on 27-Jun-2016 and are assumed to apply also for the period 31-May-2013 to 31-May-2016, which was used to calibrate the HJM model (see below)
- The Recovery Rate is assumed to be 40%
- Linear interpolation is used to approximate the CDS spreads at the half increments for which there was no market data available
- For consistency and simplicity, the discount factors used were the same as those implied from the HJM model after taking their average for each tenor (see methodology below)
- The full calculation can be found in the project repository file "*CVA/my CDS Bootstrapping v2.xlsx*"

The table below summarises the results for each of the tenors $[0.0, 1.0, \ldots, 5.0]$, with 'DF' as the discount factor, 'Lambda' as the hazard rate λ_i, 'PD' as the default probability $PD(T_i, T_{i-1}) = P(T_{i-1}) - P(T_i)$ and 'P' as the survival probability. Note that by definition PD is over a period, whereas P is cumulative.

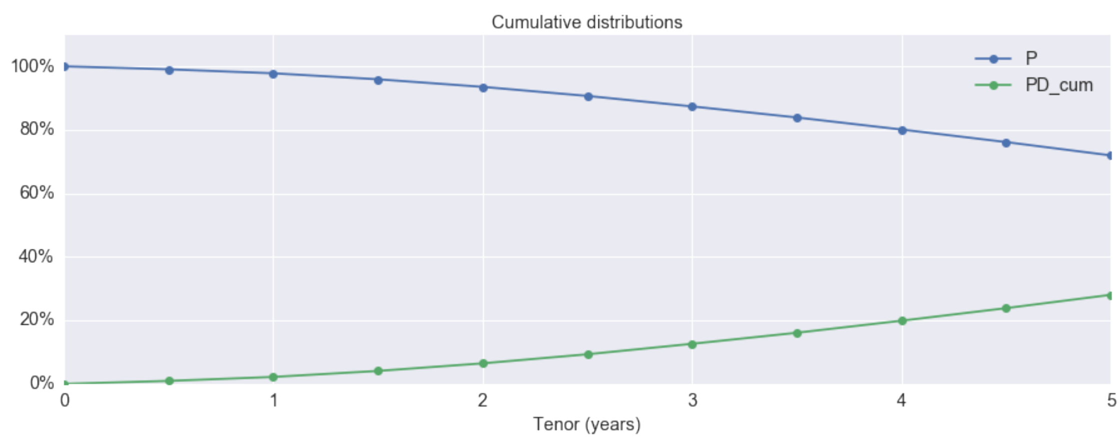
CDS Tenor	DF	Lambda	PD	P	
0.0	NaN	1.000000	nan%	nan%	100.0000%
0.5	114.400	0.995835	1.8976%	0.9443%	99.0557%
1.0	133.770	0.990963	2.2197%	1.2510%	97.8047%
1.5	167.180	0.985697	2.7798%	1.8887%	95.9161%
2.0	200.590	0.980105	3.3452%	2.3875%	93.5285%
2.5	233.965	0.974101	3.9174%	2.8579%	90.6707%
3.0	267.340	0.967832	4.4994%	3.2974%	87.3732%
3.5	296.545	0.961232	5.0170%	3.4776%	83.8957%

4.0	325.750	0.954239	5.5471%	3.7947%	80.1009%
4.5	353.200	0.946899	6.0576%	3.9605%	76.1405%
5.0	380.650	0.939187	6.5842%	4.1914%	71.9490%

The figures below show the term structure of the estimated default probability and hazard rates. As these aren't flat, this needs to be accounted in the CVA.



The plot below shows how the cumulative distribution for the survival probability P decreases with time, and vice versa for the cumulative PD since by definition this is equal to $1 - P$.



2.2 Forward LIBORs

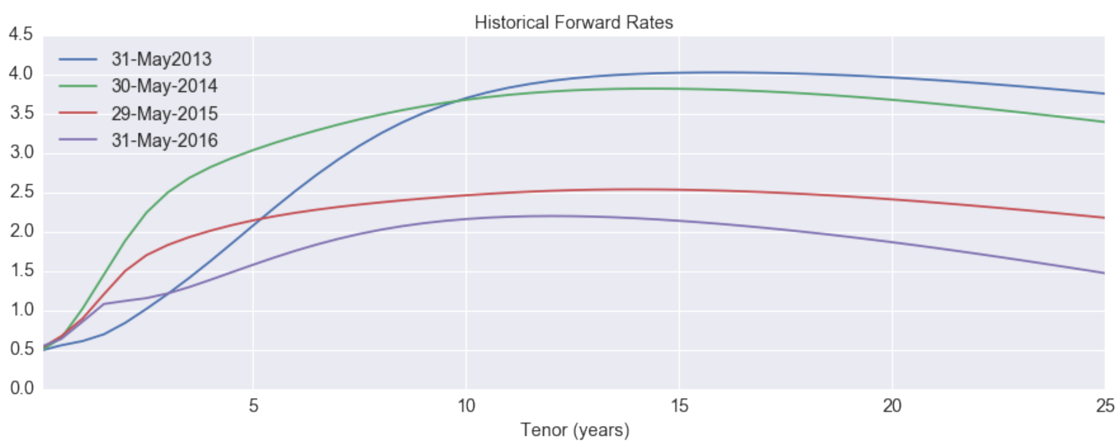
We then proceed to simulate the forward rates with the HJM model. To give a more realistic picture of the current CVA value, the HJM model was first calibrated to recent data . The steps taken are described in detail below.

2.2.1 PCA

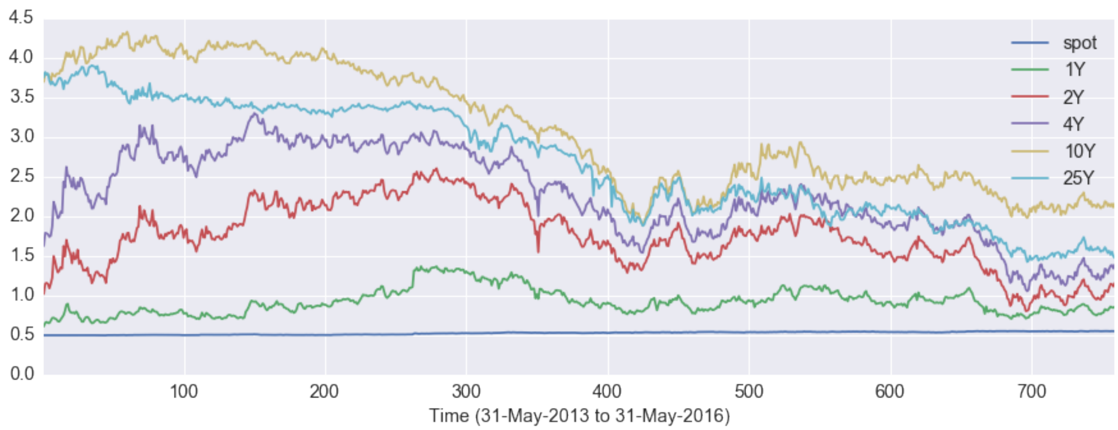
Dataset The HJM model requires a set of volatility functions which are estimated using Principal Component Analysis (PCA) on historical forward rates.

- To calibrate these functions to recent data, the forward rates were taken from the Bank of England (BoE) Bank Liability Curve (BLC) for the last 3 years (31-May-2013 to 31-May-2016). The data was taken from:
  - ukblc16_mdaily.xlsx
  - ukblc05_mdaily.xlsx
- Although we only need the short-end of the curve for the CVA calculation, we take the full curve to calibrate the HJM model, i.e. up the $25Y$ tenor. The dataset is hence constructed by taking the BLC forward curve short-end data (*"1. fwds,short end"* tab) for the tenors $[0.5Y, 1.0Y, 1.5Y, ..., 5Y]$ as it offers a better approximation to the short-end. For the remaining tenors $[5.5Y, ..., 25.0Y]$ we use the full approximation (*"2. fwd curve"* tab)
- The forward rate for the tenor $0.08Y$ in *"1. fwds,short end"* is used as a a proxy for the spot rate tenor $0.0Y$, i.e. $r(t) = f(t; t)$
- Dates with missing data values were removed
- The resulting dataset is found in the project repository under *CVA/PCA/my_ukblc_310513_310516.xlsx*

The figure below shows the forward curve for four sample dates spanning different years in the dataset. This shows how the curve can move for each of the tenors, for example, at the long-end of the curve the rates have decreased substantially in 2016 compared to 3 years ago. At the short-end of the curve we see the rates tend to increase with tenor as expected.



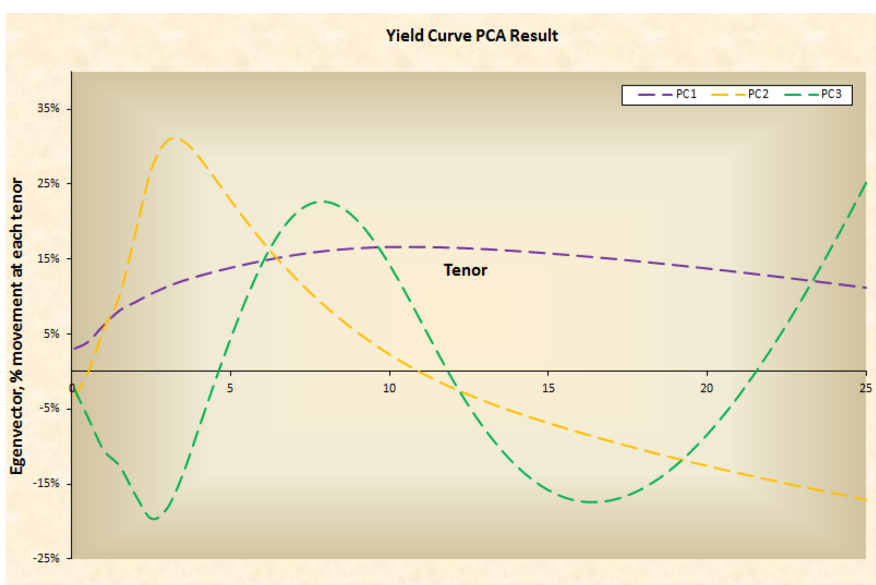
The next figure shows the historical evolution of the rate in the dataset for 3 example tenors. As mentioned, the $0.08Y$ tenor is a proxy for the spot rate. This shows the spot rate has remained quite constant throughout. The rate increases for the shorter tenors but then decreases for the longer tenors, .e.g. $25Y$ vs $10Y$ below. This could be interpreted as a lack of liquidity for the longer tenors.



Principal Components To obtain the volatility functions for the HJM model, the day-on-day changes (differences) for each tenor are calculated. This produces a set of independent random variables that can be used to get the principal components (PC), taken as the 3 largest eigenvalues and corresponding eigenvectors that explain 97.5% of the observed variance. The calculation details can be found in the project repository under "*PCA/my HJM Model - PCA.XLSM*". The results are summarised below:

	Tenor	Eigenvalue	Cum. R^2
1st largest PC	10.0	0.0076790	0.9107
2nd largest PC	3.0	0.0004034	0.9586
3rd largest PC	6.5	0.0001406	0.9752

The figure below shows the resulting eigenvectors of the principal components. In general, the largest component (PC1) is attributed to parallel shifts in the curve, the 2nd largest (PC2) to steepening/flattening (skewness) and 3rd largest to bending about specific maturity points (convexity).



We now proceed to use these eigenvectors to obtain the volatility functions for the HJM model.

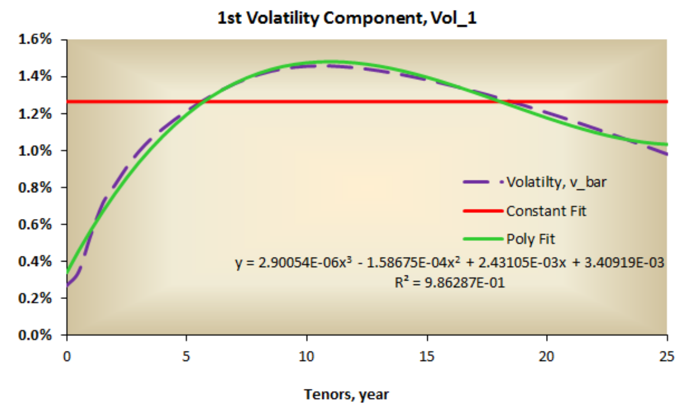
2.2.2 HJM

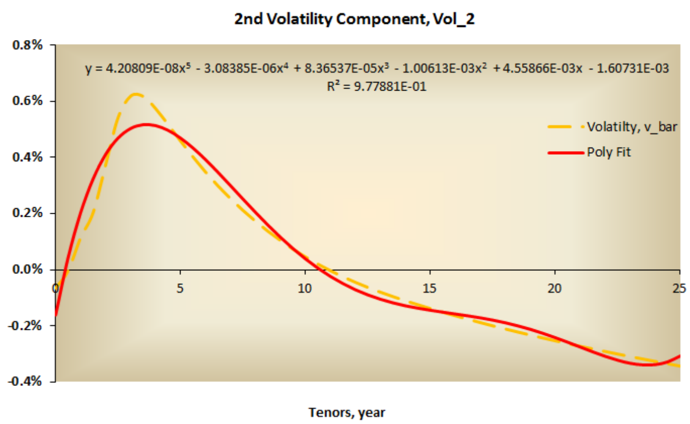
Volatility functions The volatility functions for the HJM model are defined as:

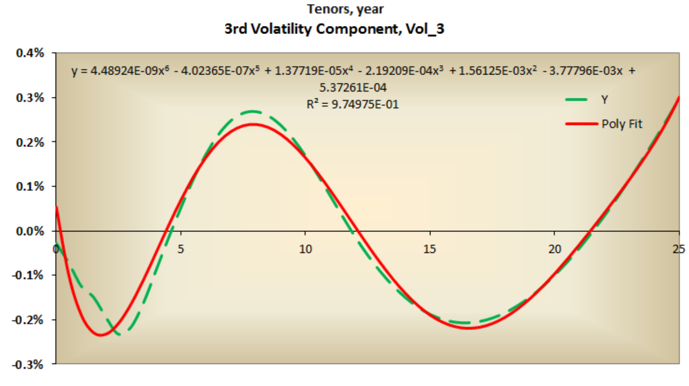
$$\mathrm{Vol}_i = \sqrt{\lambda_i}e(i) \quad \forall \quad i = 1, 2, 3 \tag{2}$$

where λ_i is the eigenvalue and $e(i)$ the eigenvector. This is equivalent to one standard deviation move in the $e(i)$ direction.

The volatility functions have to then be fitted as we need analytical functions to carry out the integration to get the drift in the HJM model. In this case, 3rd, 5th and 6th degree polynomials were fitted to guarantee a goodness of fit of over 97%. These are shown in the figures below. In particular, for Vol_1 fitting a constant (red line) is not really suited and justifies the need for a polynomial fit (however, this could also lead to overfitting issues).







Drift The drift function $\mu(t)$ is obtained by integrating over the principal components and assuming that volatility is a function of time.

Monte Carlo The calibrated volatility and drift functions above are then entered into the HJM model Monte Carlo simulation script *my_HJM_model.py*, which evolves the whole forward curve according to the stochastic differential equation (SDE):

$$d\bar{f} = \mu(t)dt + \sum_{i=1}^{3} Vol_i \phi_i \sqrt{dt} + \frac{dF}{d\tau}dt \tag{3}$$

where ϕ_i is a random number drawn from the standard normal distribution and the last term is the *Musiela correction*. For brevity the details of the model are omitted here but additional details of the MC simulation are outlined below:

- The forward curve was initialised using the the last observed forward curve (last row in the BLC data)
- Time step taken was $dt = 0.01$
- Number of simulations used was $I = 1000$
- The random variables were taken from python's numpy module *np.random.standard_normal* which draws numbers from a standard normal distribution $N(\mu = 0, \sigma = 1)$
- Due to time constrains, the antithetic variance reduction technique was not implemented in the script, so the simulation error is less negligible

To obtain an expectation of LIBOR rate in the future $L(t; T_i, T_{i+1})$, the forward rate f is selected from the corresponding tenor column $\tau = T_{i+1} - T_i$ of the HJM output, from the correct simulated time t. This is then converted to a LIBOR rate using $L = \frac{1}{\tau}(e^{f\tau} - 1)$ where $\tau_i = 0.5\ \forall\ i$ in this case.

For the IRS in question (written on 6M LIBOR for 5Y), the relevant simulated rates we need are:

L(t; T_i, T_i+1) for 6M IRS
L(t; 0, 0.5)
L(t; 0.5, 1)
L(t; 1, 1.5)
L(t; 1.5, 2)
L(t; 2, 2.5)
L(t; 2.5, 3)
L(t; 3, 3.5)
L(t; 3.5, 4)
L(t; 4, 4.5)
L(t; 4.5, 5)

The figure below shows a sample of 100 simulations for the above LIBOR rates

3 Discount Factors

The DFs are implied from the HJM Forward LIBOR for each simulation via the formula:

$$DF(0, T_{i+1}) = \prod_i \frac{1}{1 + \tau_i L(t; T_i, T_{i+1})} \tag{4}$$

which is equivalent to 'integrating under the curve'. This then gives 1000 simulations of the DF for each tenor. For simplicity, an expectation across all simulations is taken to get a single value for $DF(0, T_{i+1})$. This is then used to obtain the forward-starting discount factors as a single number $DF(T_i, T_{i+1})$ for the IRS in question, where:

$$DF(T_i, T_{i+1}) = \frac{DF(0, T_{i+1})}{DF(0, T_i)} \tag{5}$$

and by definition $DF(T_{i+1}, T_{i+1}) = 1.0$

Tenor	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
0.5	1.0	DF(0.5, 1.0)	DF(0.5, 1.5)	DF(0.5, 2.0)	DF(0.5, 2.5)	DF(0.5, 3.0)	DF(0.5, 3.5)	DF(0.5, 4.0)	DF(0.5, 4.5)	DF(0.5, 5.0)
1.0	-	1.0	DF(1.0, 1.5)	DF(1.0, 2.0)	DF(1.0, 2.5)	DF(1.0, 3.0)	DF(1.0, 3.5)	DF(1.0, 4.0)	DF(1.0, 4.5)	DF(1.0, 5.0)
1.5	-	-	1.0	DF(1.5, 2.0)	DF(1.5, 2.5)	DF(1.5, 3.0)	DF(1.5, 3.5)	DF(1.5, 4.0)	DF(1.5, 4.5)	DF(1.5, 5.0)
2.0	-	-	-	1.0	DF(2.0, 2.5)	DF(2.0, 3.0)	DF(2.0, 3.5)	DF(2.0, 4.0)	DF(2.0, 4.5)	DF(2.0, 5.0)
2.5	-	-	-	-	1.0	DF(2.5, 3.0)	DF(2.5, 3.5)	DF(2.5, 4.0)	DF(2.5, 4.5)	DF(2.5, 5.0)
3.0	-	-	-	-	-	1.0	DF(3.0, 3.5)	DF(3.0, 4.0)	DF(3.0, 4.5)	DF(3.0, 5.0)
3.5	-	-	-	-	-	-	1.0	DF(3.5, 4.0)	DF(3.5, 4.5)	DF(3.5, 5.0)
4.0	-	-	-	-	-	-	-	1.0	DF(4.0, 4.5)	DF(4.0, 5.0)
4.5	-	-	-	-	-	-	-	-	1.0	DF(4.5, 5.0)
5.0	-	-	-	-	-	-	-	-	-	1.0

The numerical values are shown below:

```
Tenor   0.0       0.5       1.0       1.5       2.0       2.5       3.0  \
0.0       1  0.995835  0.990963  0.985697  0.980105  0.974101  0.967832
0.5       0         1  0.995108  0.989819  0.984204  0.978175   0.97188
1.0       0         0         1  0.994686  0.989043  0.982984  0.976658
1.5       0         0         0         1  0.994327  0.988236  0.981876
2.0       0         0         0         0         1  0.993874  0.987478
2.5       0         0         0         0         0         1  0.993565
3.0       0         0         0         0         0         0         1
3.5       0         0         0         0         0         0         0
4.0       0         0         0         0         0         0         0
4.5       0         0         0         0         0         0         0
5.0       0         0         0         0         0         0         0

Tenor  3.5       4.0       4.5       5.0
0.0    0.961232  0.954239  0.946899  0.939187
0.5    0.965253  0.958231  0.950859  0.943115
1.0    0.969998  0.962941  0.955534  0.947752
1.5    0.975181  0.968086  0.960639  0.952816
2.0    0.980744  0.973609   0.96612  0.958252
2.5    0.986789   0.97961  0.972075  0.964158
3.0    0.993181  0.985955  0.978371  0.970403
3.5           1  0.992725  0.985088  0.977066
4.0           0         1  0.992308  0.984226
4.5           0         0         1  0.991856
5.0           0         0         0         1
```

4 Credit Valuation Adjustment

4.1 Mark-to-Market

The mark-to-market (MTM) value of the swap $V(T_i)$, i.e. the evolution of swap value over time, is obtained via:

$$
\begin{aligned}
V(T_i = 0) &= \sum_{i=1}^{11} N\tau D(0, T_i)(L_i - K) \\
V(T_i = 0.5) &= \sum_{i=2}^{11} N\tau D(0.5, T_i)(L_i - K) \\
V(T_i = 1.0) &= \sum_{i=3}^{11} N\tau D(1.0, T_i)(L_i - K) \\
&\vdots \qquad \vdots \qquad \vdots \\
V(T_i = 5.0) &= \sum_{i=11}^{11} N\tau D(5.0, T_i)(L_i - K)
\end{aligned}
$$

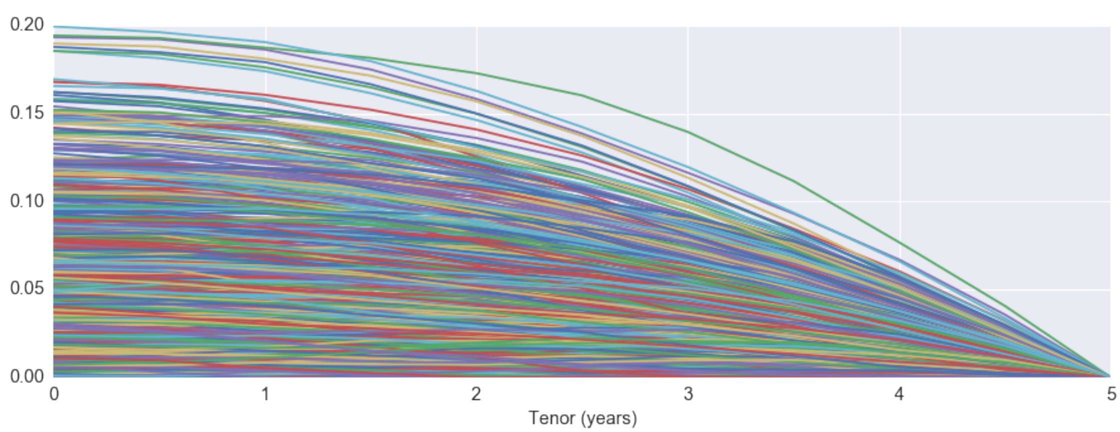
where N is the notional, K the fixed agreed rate and $L_i = L(t; T_{i-1}, T_i)$ the LIBOR effective for the period T_i. Here we assume a "par swap" (ATM) by choosing $K = L(t; 0, 0.5) \approx 0.0063845$ to have zero initial cashflows upon entering the swap.

4.2 Exposure

Finally, the exposure for each tenor E_i is calculated from the positive part of the MTM simulations as:

$$E_i = max(V_i, 0) \tag{6}$$

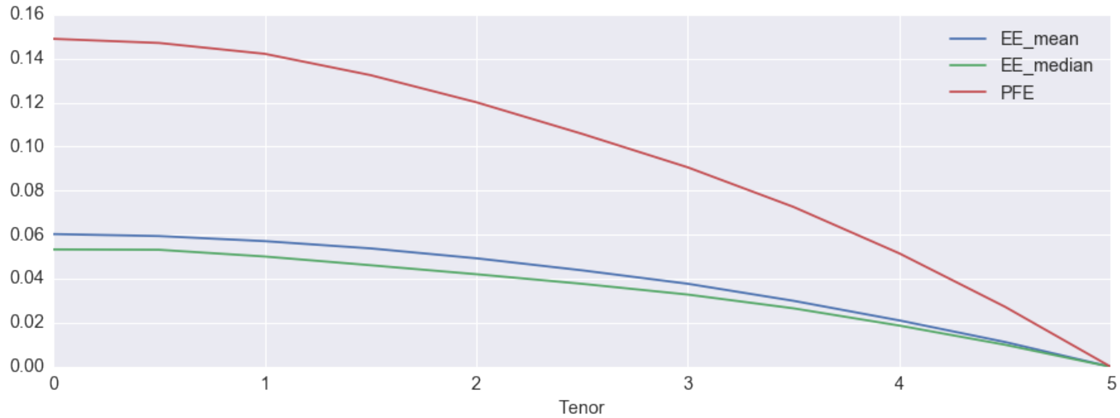
The figure below shows some sample simulations of the exposure profile.



4.3 Expected Exposure

The Expected Exposure (EE) is calculated as the median of the Exposure profile. The median is here preferred over the mean as the latter can be skewed by large observations and is more sensitive to outliers.

The figure below shows the period expected exposure EE_i as calculated from the mean and median, as well as the Potential Future Exposure (PFE), taken to be the 97.5 percentile of the E_i distribution. From this we see that using the median, the maximum EE and PFE are attained at the beginning of the swap in this case (equal to $\sim 5.33\%$ and $\sim 14.89\%$ of the notional respectively). Clearly the PFE is always bigger than the EE - in a way this tells us it is 97.5% probable that our exposure will not exceed $\sim 14.89\%$.

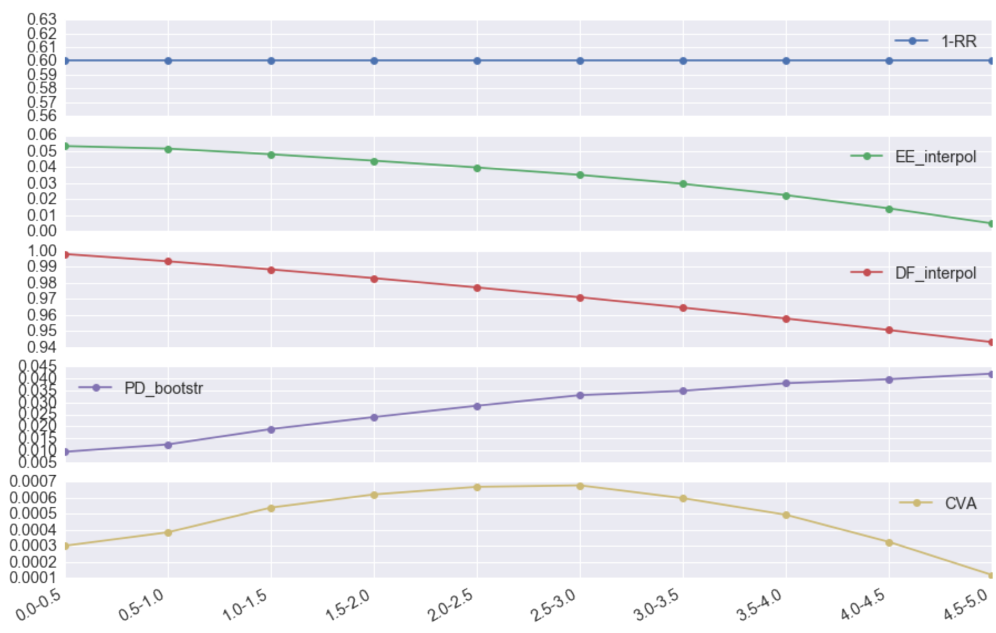


4.4 CVA

Lastly, the CVA is approximated by a linear interpolation across the tenors [1]:

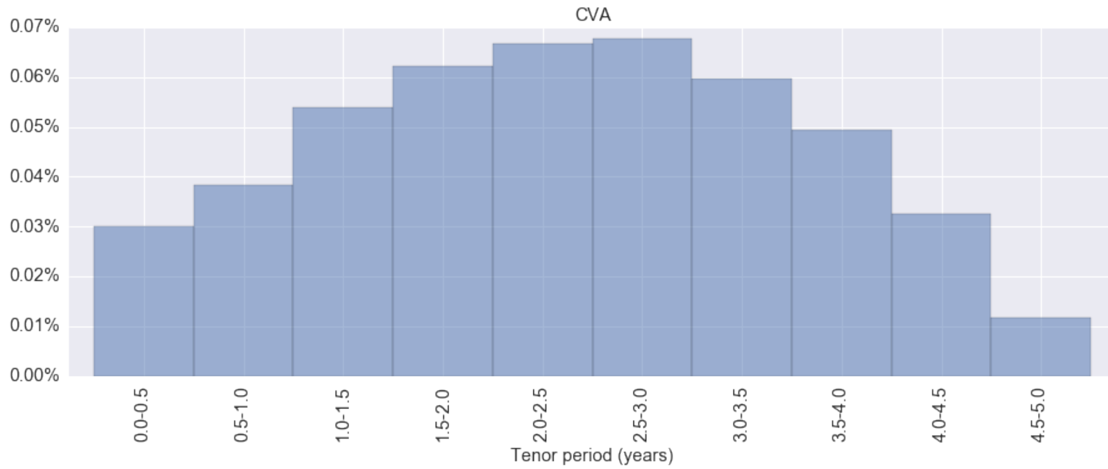
$$CVA \approx \sum_i (1-R)E(\frac{T_{i-1}-T_i}{2})DF(\frac{T_{i-1}-T_i}{2})PD(\frac{T_{i-1}-T_i}{2}) \tag{7}$$

The figure below shows each of the components in this equation, where we see their term structure isn't flat, except for the loss factor $(1-R)$.



[1]Note that the discount factors are more appropriately extrapolated using a log-linear extrapolation instead of purely linear as done here. Due to time constrains this wasn't implemented.

The figure below shows the final CVA result for each tenor range in percentage terms over the notional value. We see a 'hump'-shape, telling us the maximum is found in the $2.5 - 3.0Y$ tenor period. Based on this, the total CVA over the the lifetime of the swap amounts to 0.473% over the notional. So for example, if the notional was \$1m, the CVA would have amounted to $\sim \$4,730$, which although small is not negligible when it comes to pricing the true value of the swap.



5 Conclusion

For this hypothetical IRS scenario the CVA adjustment came out to be quite small. This was found a bit 'surprising', given AirFrance's relatively high probability of default compared to other companies CDS. Whether the CVA value arrived at is accurate or not is debatable, since through the exercise the following issues were noted:

- Typically a 'hump' structure should have been seen already in the Exposure profile, but this wasn't the case here and rather the maximum exposure was attained at the very start. This is atypical for the long-term contracts such as a $5Y$ IRS because the forward rates at these tenors are relatively high. One potential cause could be the HJM recalibration to recent data, given the rates have been historically low. Another possibility could be that the forward rates are too low compared to the estimated discount factors. On the other hand, the CVA plot did show a hump, but this is likely due to the default probability term structure. Overall this would require further investigation
- Due to time constrains, a 'shortcut' was taken when calculating the discount factors for the swap - an expectation across all simulations was taken to get a single value for each period. The effect on the valuation from this would need to be understood
- The above is in addition to all the pros and cons of the HJM model and MC associated errors. In particular, the antithetic reduction technique wasn't implemented here, yielding a higher simulation error
- The effect from accruals wasn't estimated either, although this is expected to be small

References

[1] Tanya Sandoval, *Github repository* https://github.com/tsando/CQF/tree/master/finalProject

[2] Richard Diamond, *CQF Lectures - Heath Jarrow & Morton Model*, 2016

[3] Richard Diamond, *CQF Lectures - Final Project Workshop Part I*, 2016

[4] Riaz Ahmad, *Stochastic Interest Rate Modeling*, 2016

[5] *Bank of England Yield Curves*, 2016 http://www.bankofengland.co.uk/statistics/Pages/yieldcurve/archive.aspx