CQF Final Project

CVA FOR INTEREST RATE SWAP & ARBITRAGE TRADING ON COINTEGRATION AND BACKTEST

Marcelo Pinto Brandão
MARCELOBRANDAO8@GMAIL.COM |

Table of Contents

Chapter 1 - CVA Calculation for an Interest Rate Swap

Chapter 2 - Cointegration and Arbitrage Trading on Dual-Class Shares

1. CVA Calculation for an Interest Rate Swap

How the initial selection of the historical daily forward rates period used as input in the HJM model affects its statistical accuracy and final CVA value.

1.1. Introduction

In this chapter we study the impact that the initial selection of the historical daily rates period, which is used as input in the model that projects the forward rates curve, affects the model predicted power and the CVA final value for a given contract. For this analysis the calculation is done on a Interest Rate Swap, referred as **IRS** for the rest of the chapter, and the model selected to project the forward rate curve is the Heath, Jarrow and Morton, referred as **HJM.**

The historical period chosen for this analysis are:

1) From 10/31/2015 to 10/31/2017 - Which is referred throughout this chapter as ***"recent data"*** which represents a low rate low volatility rates regime
2) From 10/01/2007 to 10/01/2009 - Which is referred throughout this chapter as ***"crisis data"*** which represents a higher rates higher volatility rates regime

The main objective is to understand how the data selection affects the goodness of fit statistic for the k-factor model ($Cum.R^2$) for the Principal Component Analysis (PCA) and how the different rates regime affects the final value of the CVA for the same derivative contract

The implementation was made in Python 3.6.3 using Jupyter Notebook, explanation of the code is not in scope of this chapter, please refer to the Python Notebook CVA_IRS-HJM.ipynb for a commented version of the code implementation

1.2. Brief Introduction of Main Concepts the CVA measure for Counterparty Credit Risk

1.2.1. CVA

Any exposure to OTC derivatives produces counterpart credit risk (CCR) which is defined as the risk that a counterpart potentially fails to meet its obligations including the non-payment of future cash flows that were agreed on derivatives contracts. The modelling of the potential future exposure that may arise from these contracts becomes a fundamental part of the risk management and need to be considered when pricing those instruments.

Credit Value Adjustment (CVA) is a correction to the fair value (or price) of derivative instruments to account for CCR. Thus, CVA is commonly viewed as the price of CCR. This price depends on counterparty credit spreads as well as on the market risk factors that drive derivatives values and, therefore, exposure.

The CVA formula is:

$$CVA = (1 - R) \int_0^T EE_t \, DF_t \, dPD_t$$

Where:

- R is the Recovery Rate
- EE_t is the Expected Exposure at time t
- DF_t is the Discount Factor at time t
- dPD_t is the default probability distribution at time t

CVA is a "charge" or "adjustment" that is applied to the price of a financial instrument so that the counterpart credit risk is taken into account:

$$\Pi^* = \Pi - CVA$$

Where:

- $\Pi*$ is the Corrected price
- Π is the fair value (or price)
- *CVA* is "charge" due to CCR

1.2.2. The Heath, Jarrow & Morton (HJM) forward rate model

HJM is a quite general framework for the modeling of interest-rate dynamics and its approach to modeling of the forward curve was a breakthrough that improved pricing and risk management in fixed income. As an alternative to deriving forward rates from modelled short-term ones, they started with a model for the whole forward rate curve.

It uses real historical data of the movement of the forward rates curve and incorporate them into the pricing methodology, which makes the fitting to a market yield curve a natural process within the HJM. Starting with the forward rate equation $f(t, T) = -\frac{\partial}{\partial T} lnZ(t,T)$ and using the bond price "GBM" SDE to derive the evolution of forward rates, we come at:

$$dF(t;T) = \frac{\partial}{\partial T}\left(\frac{1}{2}\sigma^2(t,T) - \mu(t,T)\right)dt - \frac{\partial}{\partial T}\sigma(t,T)dX$$

Which is the Stochastic Differential Equation (SDE) for a forward curve. The SDE carries the drift of the bond price $\mu(t, T)$.

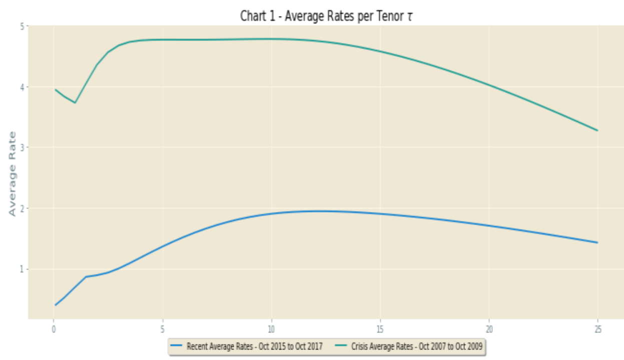
1.3. HJM Model Implementation and Analysis of Results

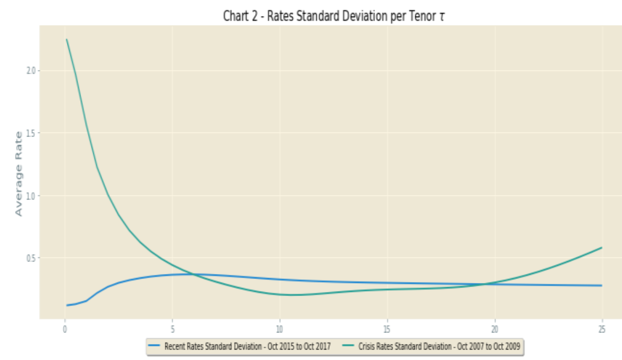
1.3.1. Data Preparation

For this implementation of the Heath, Jarrow & Morton forward rate model we use as initial data the daily forward rates from the Government Liability Curve (GLC) data provided by the Bank of England found at https://www.bankofengland.co.uk/statistics/yield-curves. For this data set we have daily rates beginning at January 04, 2005 until October 31, 2017.

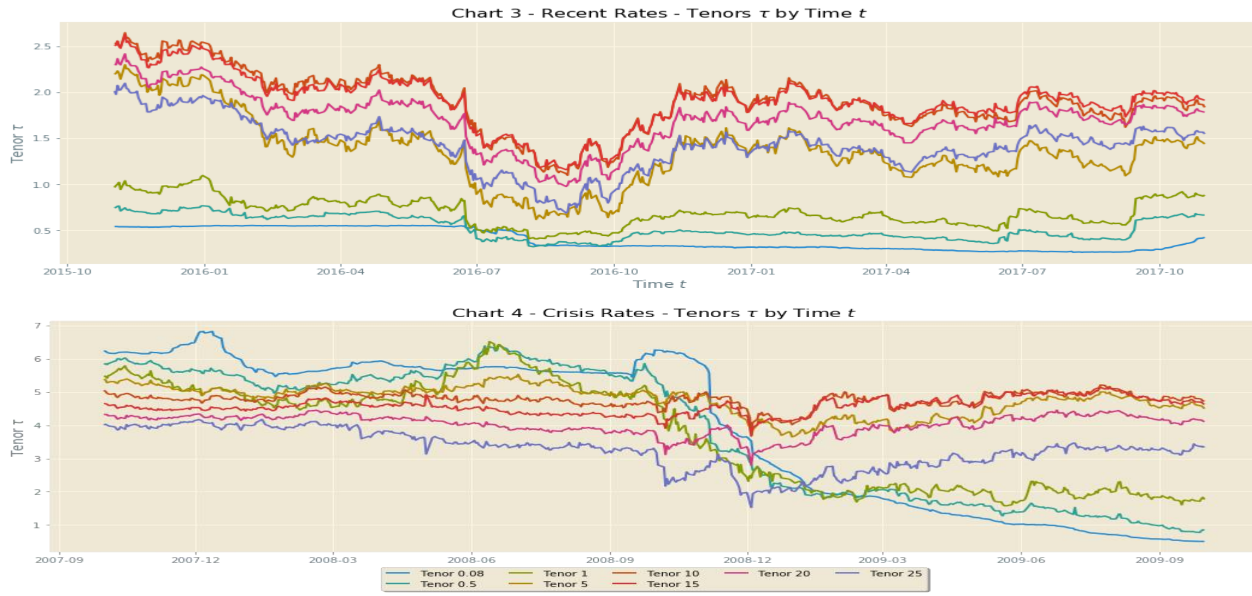
The available tenors for the set are: 0.08 0.5 1.0 1.5 2.0 2.5 3.0 ... 22.0 22.5 23.0 23.5 24.0 24.5 25.0

Chart 1 and **Chart 2** represents the periods Average Rates and Standard Deviations, it is possible to visualize the 2 different rates regimes, most recent data average rates reflects the low rates regimes from the last years with a stable standard deviations, while during the financial crisis rates were at higher levels and with higher volatilities especially in the lower and higher ends of the curve.





A graphic representation of the evolution of selected Tenors(τ) (0.08, 0.5, 1, 5, 10, 15, 20 and 25) by the time is presented bellow on charts 3 and 4. Here it is also possible to visualize the different behaviors of the tenors on each rate regime. Short term rates (0.08 and 0.5) presents a more stable development on the last 2 years when compared to the crisis period.



1.3.2. Covariance Matrix Estimation

To estimate the covariance matrix, we need to obtain daily differences rates for each tenor. For low rates regime it is best to conduct the Principal Component Analysis (PCA) on log-differences in forward rates: $log\ f(t+\Delta t) - log f(t)$ for each tenor. We applied log differences for the most recent data series and differences for the crises data series. Chart 5 and 6 shows the results.

1.3.3. Covariance Matrix Decomposition

Σ is the re-annualized covariance matrix of forward rates changes. Any symmetric matrix can be decomposed according to the spectral theorem:

$$\Sigma = V \wedge V'$$

Where :

∧V == Diagonalised Vectorized matrix matrix with **eigenvalues** of **eigenvectors**

For each eigenvector e(i), the first entry is the movement of the tenor 0.08, the second entry is related to the 0.5 Tenor and so on. Our PCA will focus on the 3 main eigenvectors which are selected by the corresponding 3 largest eigenvalues.

1.3.4. Factor Attribution

Parallel shift in overall level of rates is the largest principal component of forward curve movement, common to all tenors.

Steepening/flattening of the curve is the second important component. Inverted curve would have a different shape for the PC2.

Bending about specific maturity points is the third component to curve movement that mostly affects curvature (convexity).

Charts 7 and 8 represents the 3 PCA for each covariance matrix.

Chart 7 - Principal components for Recent Rates

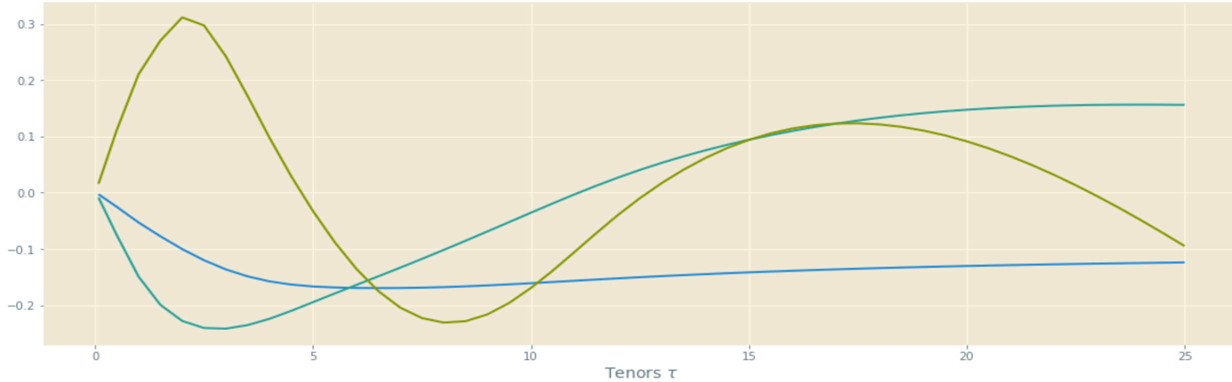
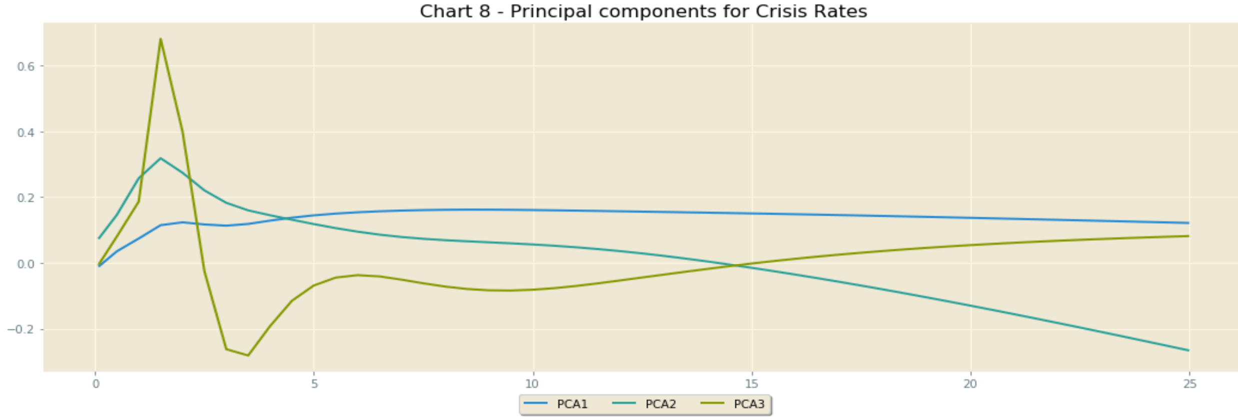


Chart 8 - Principal components for Crisis Rates



Principal eigenvalues recent data:

[2.13769377e-03 1.07383656e-04 7.34959550e-05]

The 3 PCA explains: 96.89 % of variation in the BLC curve

Principal eigenvalues crisis data:

[0.00342655 0.00111086 0.00080282]

The 3 PCA explains: 78.75 % of variation in the BLC curve

Comparing both cumulative goodness of fit statistic for the k-factor model ($Cum.R^2$) we can conclude that the eigen-decomposition based on the recent data has a stronger predictive power. This is since the PCA analysis using the crisis dataset produces a higher uncertain attribution caused by a mix of higher rate period (Chart 4 - 2007/09 to 2008/12) and a regime of lower rates (Chart 4 - 2008/12 to the end). Even though the Crisis dataset produces a poor PCA Component attribution we are going to continue the exercise in order to understand its impact on the final CVA calculation for the IRS.

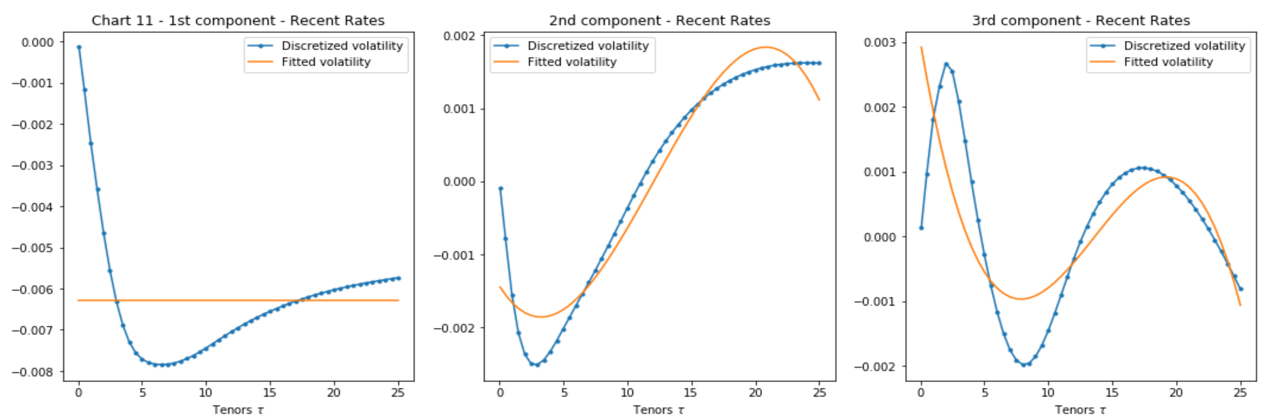
1.3.5. Obtain Volatilities Functions

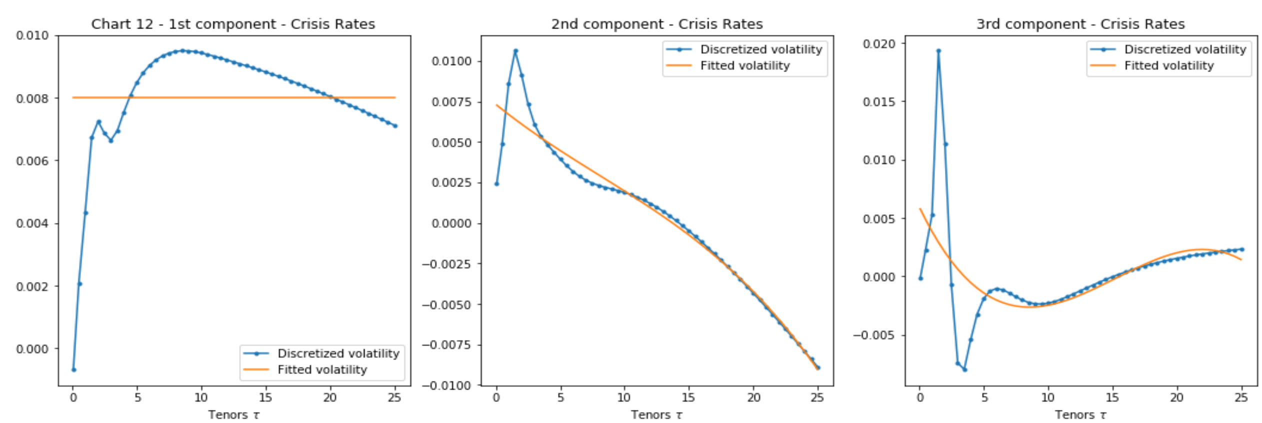
Through calibration bc **PCA** and **fitting by cubic spline** the volatilities functions $Vi(t, \tau)$ and therefore, drift function $m-(t, \tau)$ are obtained. Both are functions of Tenors τ. HJM SDE applies at any point fj of the forward curve. We have a vector stochastic process.

$$f(t, T) \Rightarrow f-(t, T-t) = f-(t, \tau)$$

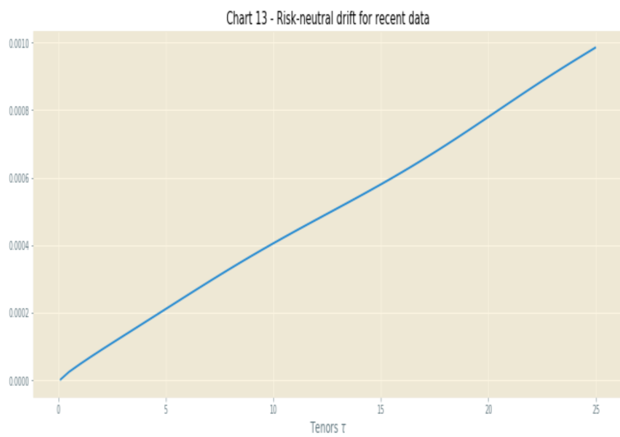
can be re-parametrized for tenor time $\tau_j = T_j - t$, index $j = 1...N$ refers to tenor 0.5Y, 1Y, ... etc.

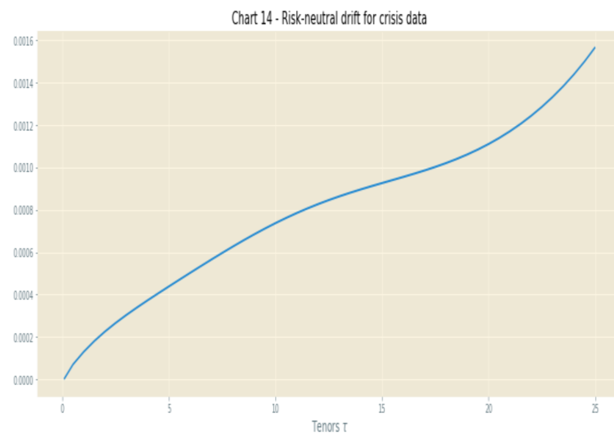
Volatility Fitting Fit Volatility Functions $\bar{v}(t, \tau)$ from discretized versions.





Drift calculation Drift $m(t, \tau)$ is calculated using numerical integration over fitted volatility functions (No Mursiela parameterisation for now)





1.3.6. Monte Carlo Simulation

At this point we have all necessary elements to execute the Monte Carlo Simulation which will provide a set of forward rates which will be used to calculate the discount factors that will be used as input on the IRS price calculation and on the CVA charge of this contract.

We simulate the instantaneous forward for τ_j for all tenors.

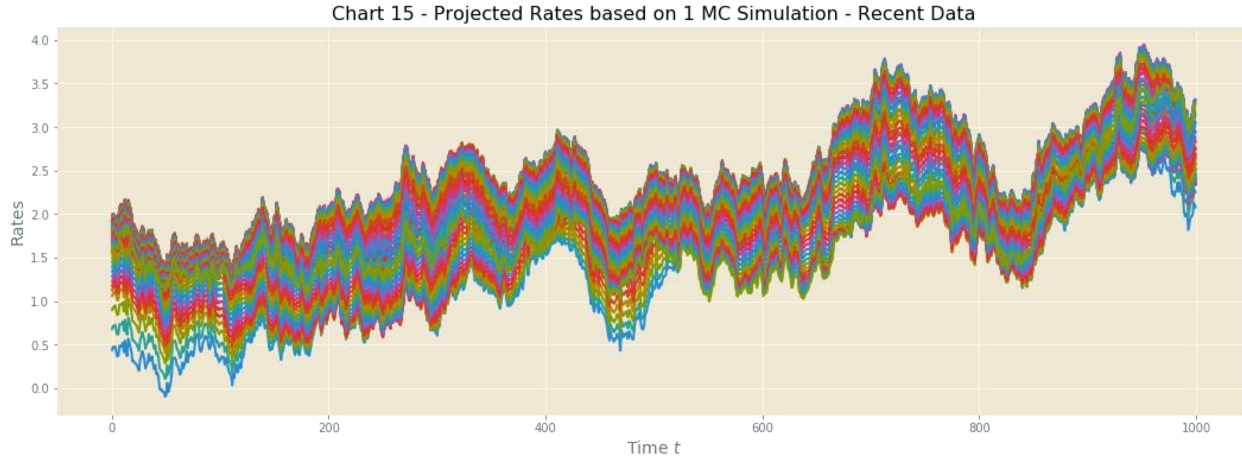
$$\bar{f}_j(t+dt) = \bar{f}_j(t) + d\bar{f}_j$$

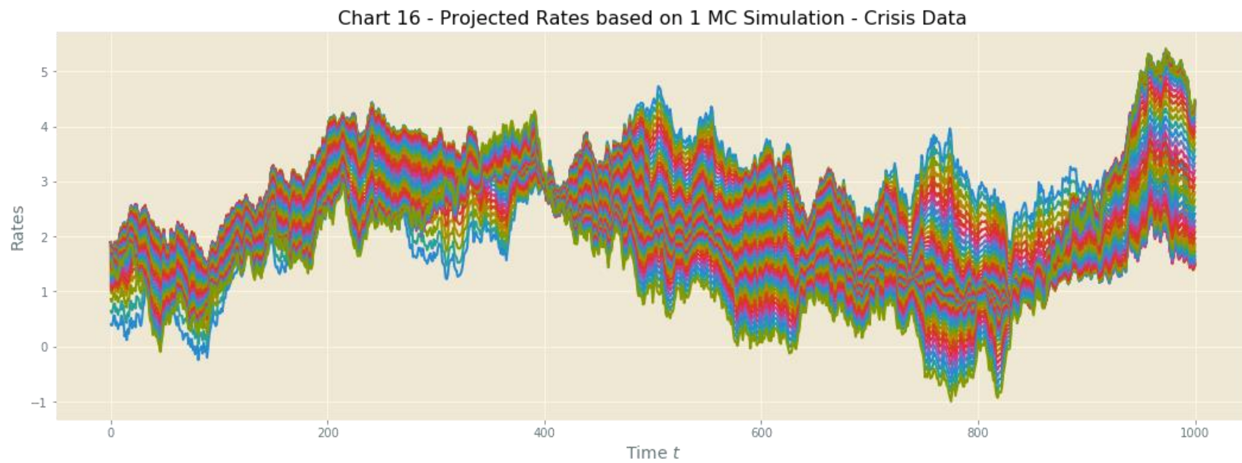
Write the multi-factor HJM SDE (Musiela Parametrization for τ)

$$d\bar{f}_j(t,\tau) = \bar{m}(\tau)dt + \sum_{i=1}^{k=3} \bar{\nu}_i(t,\tau)dX_i + \frac{\bar{f}(t,\tau)}{\partial\tau}dt$$

$$\bar{\nu}_i(t,\tau)dXi = \sqrt{\lambda_i} e^i_\tau \, \Phi_i \, \sqrt{\tau t}$$

The uncertainty around the k N linearly independent factors is simulated by the uncorrelated Brownian Motions dXi times volatility





1.4.Interest Rate Swap Contract Pricing

For this exercise the Interest Rate Swap (IRS) Contract Specifications are:

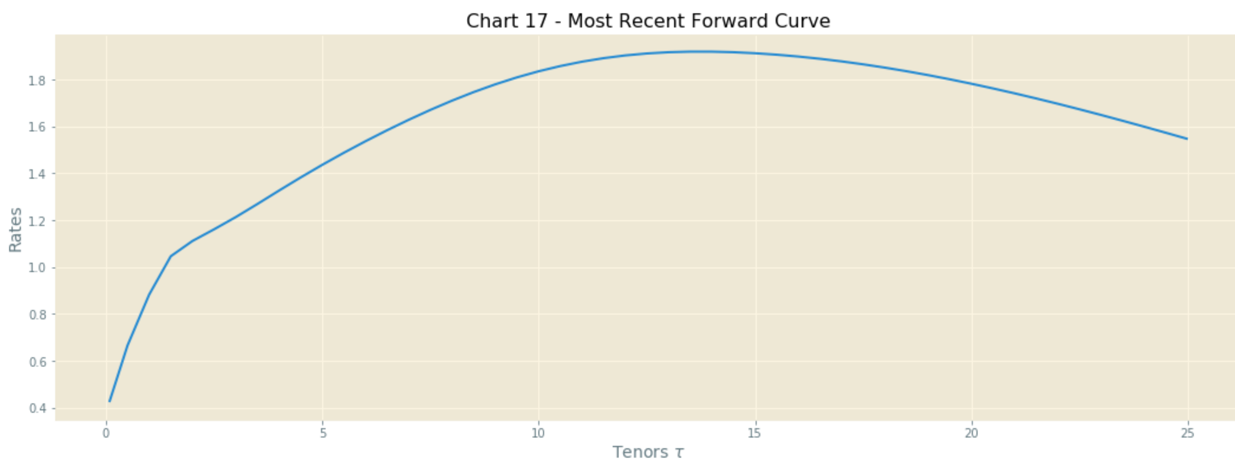
pay_leg = 0.01425

maturity = 5

dt = 0.5

notional = 1000000

Since to calculate the Fair Value of the IRS we are using the last forward rate from the initial database. The price for IRS does not depend on the initial period of choice. For the most recent forward curve (chart 17) we can calculate the IRS Fair Value.



IRS Present Value is USD: 178.03

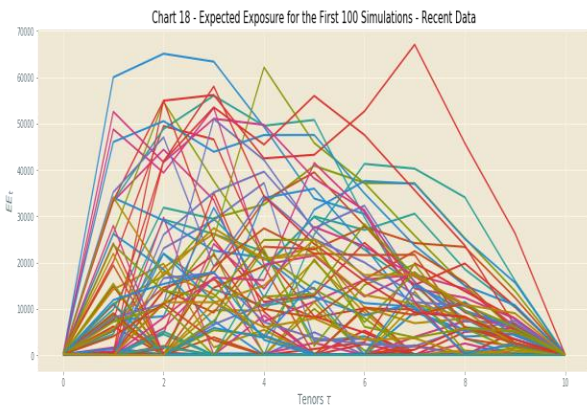
1.1. CVA Calculation for IRS

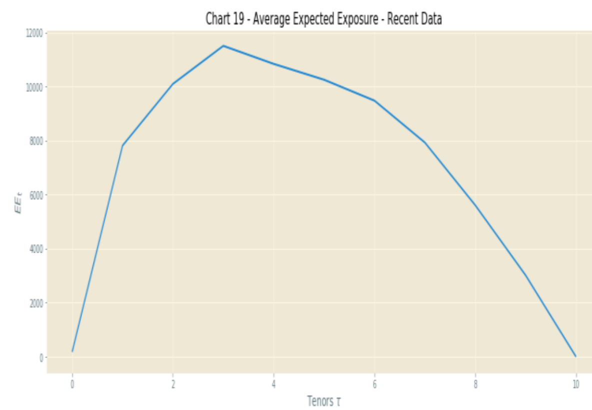
1.1.1. Interest Rate Swap Expected Exposure

At this point are ready to use the HJM output for CVA calculation, the following example is based on 2 Monte Carlo runs of 1000 simulations each (one for recent data period and other for the crisis period). Coming back to the CVA formula: $CVA = (1 - R)\int_0^T EE_t\, DF_t\, dPD_t$

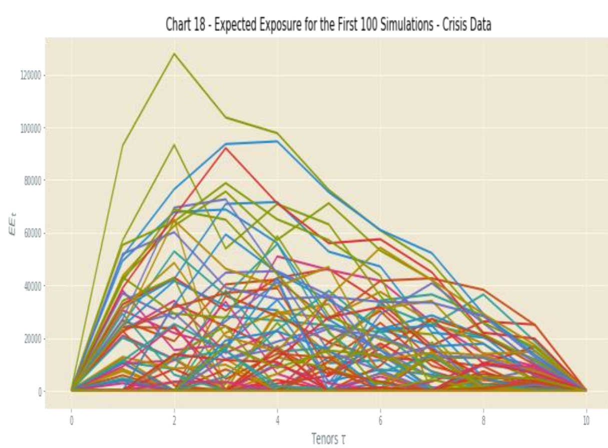
EE_t is the Expected Exposure at time t where $EE_t = E^Q[max(MTM,0]$, and we calculate it as average exposure for the 1000 simulations per Tenor τ.

For the recent data the results are:





For the crisis data the results are:



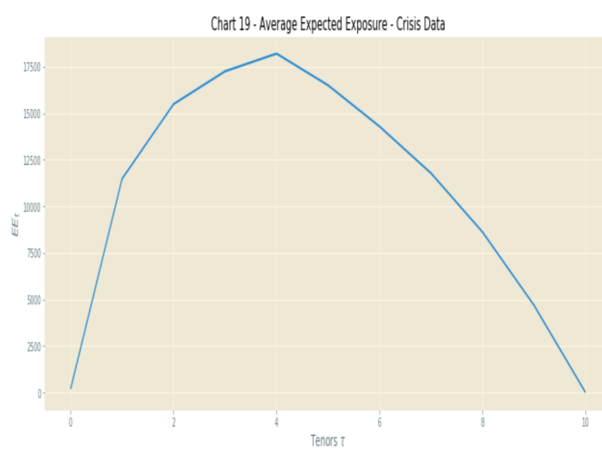


Table 1 - Recent Rates Average Expected Exposure :

	0	1	2	3	4	5	6	7	8	9	10
0	178.0	7817.0	10102.0	11510.0	10843.0	10256.0	9482.0	7932.0	5610.0	3015.0	0.0

Table 2 - Crisis Rates Average Expected Exposure :

	0	1	2	3	4	5	6	7	8	9	10
0	178.0	11475.0	15489.0	17252.0	18204.0	16505.0	14291.0	11779.0	8621.0	4687.0	0.0

Average Expected Exposure Crisis/Recent (%) :

	0	1	2	3	4	5	6	7	8	9	10
0	0.0	46.8	53.32	49.88	67.89	60.93	50.71	48.5	53.65	55.44	NaN

Comparing the 2 situations it is possible to see how the more volatile period for forward rates during the 2008 financial crisis reflects on 45% to 50% higher Expected Exposures EE_t for the 2008 period.

1.1.2. Loss Given Default (LGD), Discount Factors (DFt) and Default Probability (dPDt)

For the following example we will be assuming:

lambda_rate=0.05

recovery_rate = 0.4

We will vectorize all variables in order to facilitate final calculations. For Loss Given Default $LGD = (1 - Recovery\ Rate)$ in vector format:

	0	1	2	3	4	5	6	7	8	9
0	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6

Discount factors DF_t are extracted from the most recent forward curve, we use the average between 2 discount factors:

	0	1	2	3	4	5	6	7	8	9
0	0.998342	0.993961	0.987857	0.981283	0.974786	0.967932	0.960546	0.952616	0.944219	0.935448

Default Probability formulas is $dPD_t = PD(t_{i-1}, t_i) = P(0, t_{i-1}) - P(0, t_i)$, in vector format :

	0	1	2	3	4	5	6	7	8	9
0	0.02469	0.02408	0.023486	0.022906	0.022341	0.021789	0.021251	0.020726	0.020215	0.019715

For Expected Exposure EE_t we use the difference of EE_t between 2 consecutive tenors :

Exposure Array for Recent Data :

	0	1	2	3	4	5	6	7	8	9
0	7638.570581	2285.65158	1408.140856	-667.536589	-586.688493	-773.937204	-1550.261005	-2321.634344	-2595.069314	-3015.261565

Exposure Array for Recent Data :

	0	1	2	3	4	5	6	7	8	9
0	11297.024803	4013.590328	1763.130605	952.233268	-1698.515638	-2214.628959	-2512.021232	-3158.323875	-3933.503329	-4687.01147

Using the $CVA = (1 - R)\int_0^T EE_t\, DF_t\, dPD_t$ formula on with the arrays :

IRS Present Value is USD: 178.03 CVA

charge for Recent Dataset is: 29.36

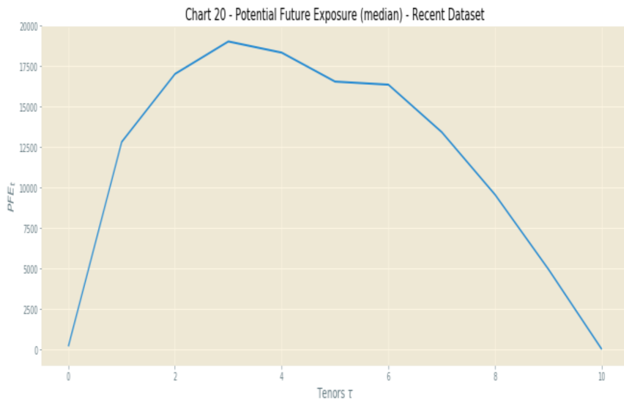
CVA charge for Crisis Dataset is: 46.79

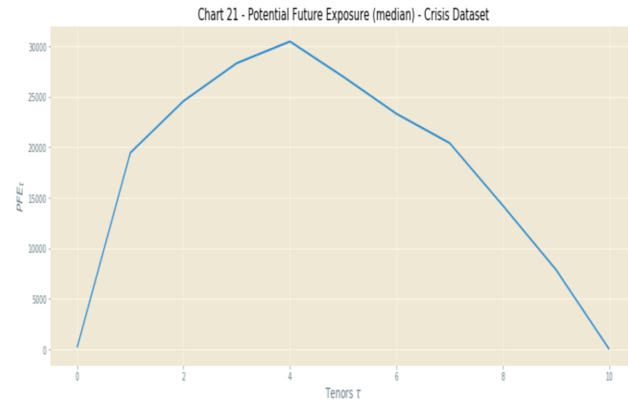
IRS Final Price is for Recent DataSet is: 148.67

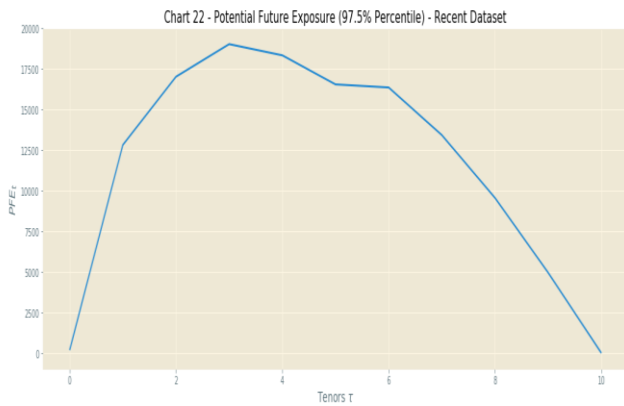
IRS Final Price is for Crisis DataSet is: 131.24

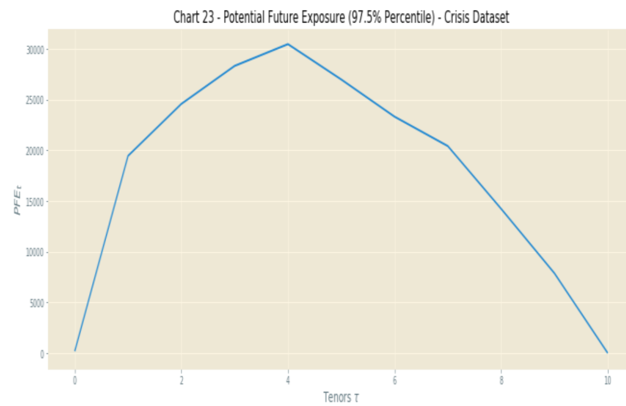
CVA Charge using the 2008 Financial Crisis Forward Rate data as the initial input into the HJM

Additionally, using the output from the Monte Carlo Simulation, we can produce **Potential Future Exposures profiles (PFE)**, the following charts demonstrate PFE for both cases calculated as: Median of Positive Exposure (Chart 20 and 21) and as the 97.5% percentile (charts 22 and 23).









1.1.3. Conclusion

Comparing both results the turbulent and highly volatile rates regime of the financial crisis reflects in a considerably higher (52%) CVA when compared with the recent rates regime implementation but as expected it produces a lower quality principal components analysis, ,$Cum.R^2$ for the second case is 78.75% compared with a 96.98% fit for the recent data case, given the unclear attribution which can be observed in chart 12 (PCA components do not match their roles (parallel shift, skew and twist) of the curve).

But since that the effect of a different rate regime than the one that we see nowadays have relevant impact on the CVA calculation it would be advisable to account for the stressed periods on the credit calculation with the inception of a different and more volatile rate regime on the models. Maybe by averaging a current CVA calculation with recent regimes with a CVA from a stressed period in which we can get a good stat fit.

2. Cointegration and Arbitrage Trading on Dual-Class Shares

Cointegration and mean reverse strategy applied to different shares classes from the same bank, Itaú Unibanco in Brazil.

2.1 Introduction

Pair trading is a popular statistical arbitrage strategy, it consists in trading the spread (long/short position) of shares, which usually possess the same "economic" factors, when a dislocation between the two prices paths is observed. To set up such strategy we split the implementation in 2 parts: the selection part (which pairs combination provide the right conditions (stationary state)) and the implementation of the trading strategy with back-test so an analyze on cumulative P&L behavior is possible.

In this chapter we introduce the concept of stationarity of a linear combination of assets, leading to cointegration and unit root tests to determine if a pair of Brazilian bank stocks of different classes for the same bank (ITUB4 and ITUB3) an initial basic strategy is developed on the first step and results presented based on a developed back-test algorithm. On a second step a couple of rules are introduced with the objective of try to improve the results of the basic strategy.

The implementation was made in Python 3.6.3 using Jupyter Notebook, explanation of the code is not in scope of this chapter, please refer to the Python Notebook BanksArbitrage.ipynb for a commented version of the code implementation

2.2 Dual Class Shares in Brazil - A Background

In the Brazilian financial markets, a considerable number of companies have 2 classes of shares traded on the exchange: **Ordinary shares (ON)** and **preference shares (PN)**:

Common Shares - Gives its holder the right to vote in shareholders' meetings and participate in the economic profit of the company. It is represented by the number three at the end of the share abbreviation, in our example: ITUB3

Preference Shares[1] - Grants its holder preference in the receipt of dividends and, in the case of liquidation, preference in capital reimbursement. It does not give the holder the right to vote in shareholders meetings. They are represented by the number four, five, six, seven or eight on the stock code, in our example: ITUB4

As described on *An Assessment of Dual-Class Shares in Brazil: Evidence from the Novo Mercado Reform by Matt Orsagh,* historically, Brazilian companies have had a dual-class structure with common shares (voting) and preferred shares (nonvoting). To compensate for no voting rights, preferred shares pay (in theory) higher dividends or have "tag-along" rights (takeout rights on a sale of control). This enables holders of voting shares to control companies by owning less of the company's total equity, which is a deviation from the proportionality principle ("one-share one-vote").

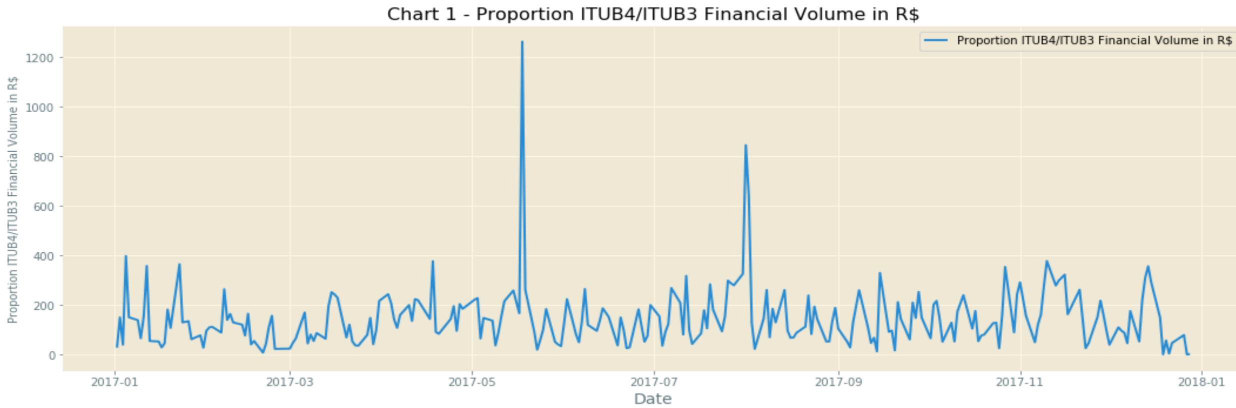
The preference shares have been introduced in Brazilian markets around mid-1970 in an attempt to incentivize more companies to open its capital on the exchange without giving up control, the initial law of 1976 allowed companies to issue 2/3 of total shares available as preference shares. This instrument was widely used by companies throughout the 80s, 90s until beginning of years 2000 and it PN shares quickly became the most liquid class and most of the times the only option for big institutional investors.

By earlier 2000 with a more mature market a process to decrease the size of preference shares was initiated with a change on the 1976 law decreasing the PN proportion to 50/50 on the subsequent years new rules were introduced with the objective to extinguish the market of PNs especially the rules about governance levels which states that for a company to be included on the exchange highest level of governance they should not have any PN traded on the market. The companies that inherited this characteristic are exchanging PN shares for ON, but there are still a lot of companies that possess both classes, especially on the banking industry.

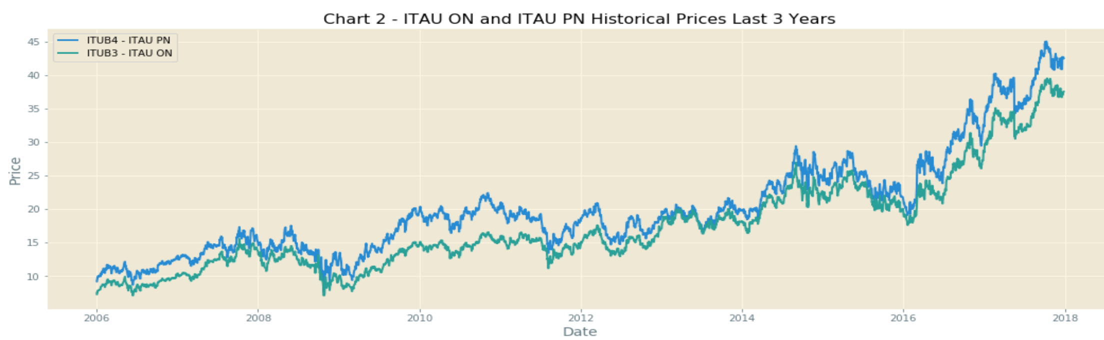
On the example on this chapter our choice was to use the 2 classes from Itaú Unibanco *(https://www.itau.com.br/relacoes-com-investidores/o-itau-unibanco/sobre-o-itau-unibanco),* the bank is the result of the merger of Banco Itaú and Unibanco, which occurred on November 4, 2008 to form Itaú Unibanco Holding S.A, the largest financial conglomerate in the Southern Hemisphere and is the 10th largest bank in the world by market value. The bank its listed at the B3 in Sã Paulo and in NYSE in New York. It currently is the biggest Latin American bank by assets and market capitalization.

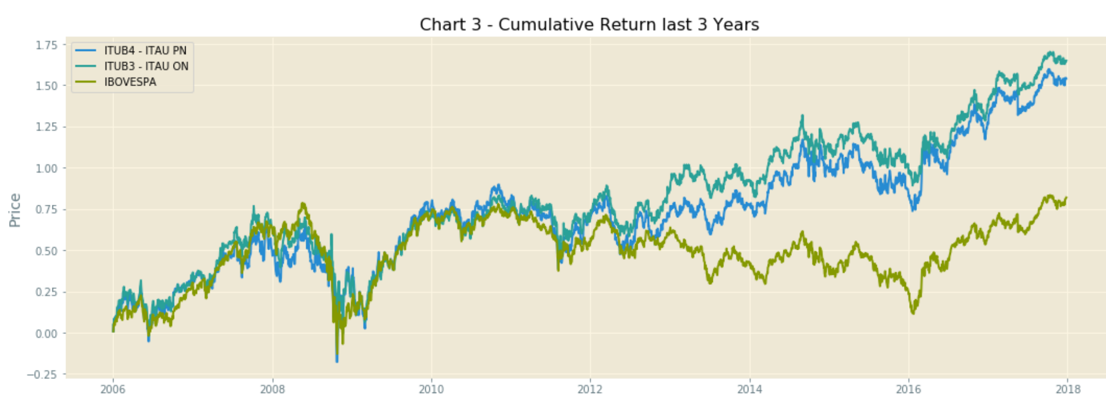
Although in theory the differences of rights on the share classes would justify the premium/discount that are observed in the market in practice on Itaú Unibanco case it does not happen since dividends and tag-along rights are the same for both classes, still the preference class have historically traded over a premium in relation to the common type, the two factors that are left to explain the premium/discount relationship are value of shareholders rights and the value of liquidity of the shares.

The common accepted explanation around the premium of ITUB4 over ITUB3 is in relation to the better liquidity of ITUB4 (chart 1), as an example on average in 2017 the daily financial volume traded on ITUB4 is 145 times the volume in ITUB3 in Brazilian Reais (R$).



As expected those shares hold a strong positive relationship our first objective is to understand if a linear combination between these 2 stocks produces a stationary time series that can be translated in a opportunity for a mean-reverting strategy. The second objective is to try to improve the strategy's result by optimizing the strategy construction methodology.





Charts 2 and 3 above demonstrated how the 2 share classes move together for the last 10 years with the spread between the 2 widening a lot between 2009 and 2013 and more recently after 2016. It is going to be interesting to see how the returns for the strategy behaves on those 2 periods. Scatter plots (chart 4 and 5) of the 2 stocks suggest a strong positive correlation, but it is not a guarantee of stationarity between the time series, but if we take a linear combination of these series it can sometimes lead to a stationary series. In this case we say that the series are **cointegrated.** The mathematical definition: Let x_t and y_t be two stationary time series, with a, b as constants. If the resulting series of $ax_t + by_t$ is stationary, then we say that x_t and y_t are cointegrated.

Chart 4 - ITUB4 x ITUB3 Scatter Plot

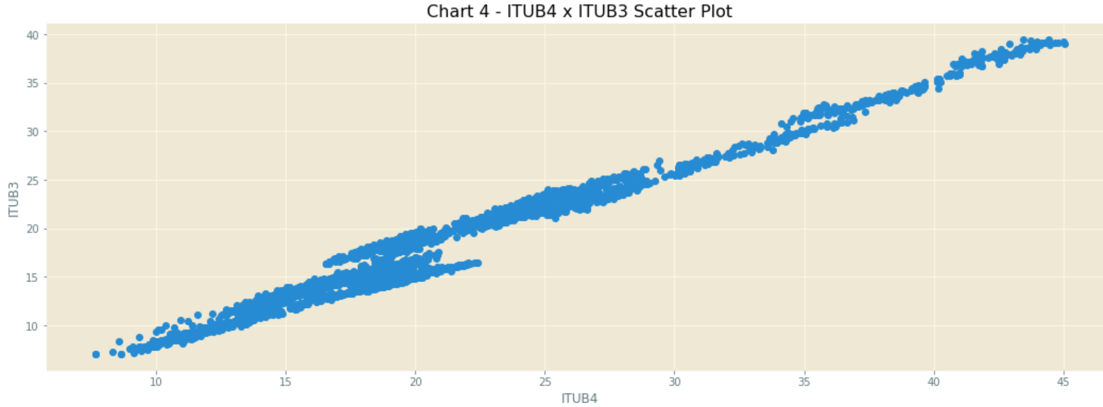
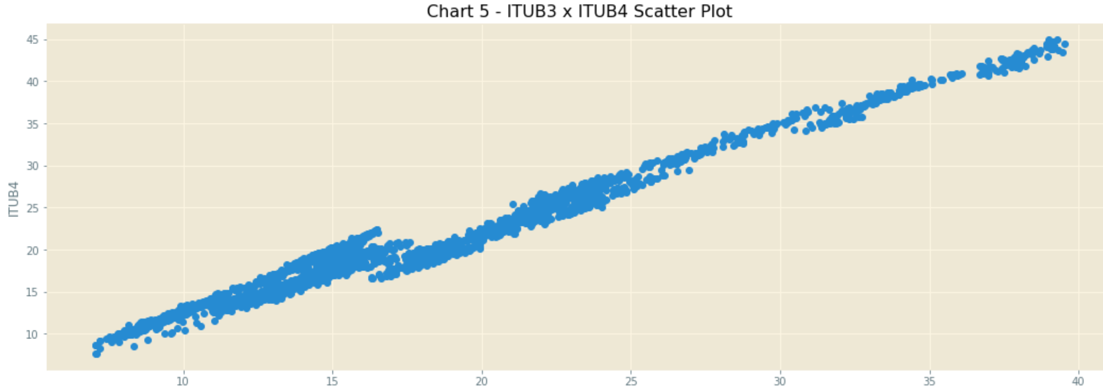


Chart 5 - ITUB3 x ITUB4 Scatter Plot



2.3 Cointegration Testing for Stocks Pair (ITUB4/ITUB3)

2.3.1 The Engle-Granger Implementation Steps:

Step 1 - Estimate the regression: $y_t = \alpha + \beta x_t + e_t$ (1), where:

y_t is the dependent variable time series of prices

α is the intercept of the regression

β is the multiplier of the independent variable, refereed on this chapter as Hedge Ratio

e_t is the residuals (or error term) of the regression

In order to determine which is the most appropriate relationship between the stocks is important to test both stocks as dependent and independent variables, which leads us to 2 regressions for the example:

$$ITUB4 = \alpha + \beta IUTB3 + e_t \quad (2)$$

$$ITUB3 = \alpha + \beta ITUB4 + e_t \quad (3)$$

We consider $\alpha = 0$(*the intercept*) the statistical reason behind this is that the linear regression model fits 2 variables one is hedge ration and the other the intercept, by setting the intercept to 0, the model will only fit the hedge ratio which reduces "overfitting".

Step 2 - Calculate the residuals (error term) e_t. That is, given that we consider $\alpha = 0$ in (1), we can write the residuals as $e_t = y_t - \beta x_t$ (4), for the example the residuals would become:

$$e_{t1} = ITUB4 - \beta ITUB3$$

$$e_{t1} = ITUB3 - \beta ITUB4$$

Step 3 - Perform a unit root test on e_t to determine if it is I(0). The augmented Dickey-Fuller (CADF) test for stationarity on residuals (4) provides the unit root test required, the hypothesis H0 is that **the prices series are not stationary**. The test must less than the critical value for H0 to be rejected and to conclude that the time series is stationary and not a random walk.

2.3.2 ADF Test - Mathematical background

The general set up for a CADF Test is:

$$\Delta Z_t = \beta_0 + \beta_1 t + \delta Z_{t-1} + \sum_{i=1}^{k} \phi_i \Delta Z_{t-1} + \epsilon_t$$

Where *k* represents the lag in the time series and BIC or AIC can be used to get the optimal lag k, for the example in this chapter the model with lag order 1 is used, additionally the ADF Statistics varies depending on whether an intercept term (*β*) are included, as explant on the The Engle-Granger Implementation Steps the spread should have zero mean and we set the intercept term to 0, but it is noted that some practitioners tend to keep the (*β*). For our implementation the ADF formulation used is:

$$\Delta Z_t = \delta Z_{t-1} + \phi \Delta Z_{t-1} + \epsilon_t$$

The null hypothesis (H0) of the ADF test is that the spread δZ_{t-1} is a unit-root process, and the alternative (H1) is that the process is a stationary process:

H0 : $\delta = 1$

H1 : $\delta < 1$

The ADF Statistic is defined as the T-Statistic for the δ coefficient. For this example, the ready implementation of the test present on the **statsmodels.tsa** package in Python is used and the **critical values** for the test statistic at the 1 %, 5 %, and 10 % levels are based on Based on MacKinnon (2010):

- 1pct = 3.51%

- 5pct = 2.89%
- 10pct = 2.58%

As an initial example we are going to implement the strategy and back test for the entire data set based on the Engle-Granger results for the first year of data (252 business days). Meaning that we are going to establish a fixed long/short relationship and keep it fixed for the entire time range (10 years).

This is recognizable not an ideal approach but will serve well as an introductory example.

Table below shows the results for the regression of which pairwise relationship possible (2), one important point is that the CADF Test performed using the Python's package statsmodels.tsa.stattools produces the result based on the optimal lag selection *

* 'AIC' the number of lags is chosen to minimize the corresponding information criterion

Regression Results for ITUB4 as dependent variable:

```
Regression Results for ITUB4 as dependent variable :

                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       1.000
Model:                            OLS   Adj. R-squared:                  1.000
Method:                 Least Squares   F-statistic:                 8.547e+05
Date:                Sun, 07 Jan 2018   Prob (F-statistic):               0.00
Time:                        19:02:16   Log-Likelihood:                 63.119
No. Observations:                 252   AIC:                            -124.2
Df Residuals:                     251   BIC:                            -120.7
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
x1             1.2479      0.001    924.512      0.000       1.245       1.251
==============================================================================
Omnibus:                        2.605   Durbin-Watson:                   0.679
Prob(Omnibus):                  0.272   Jarque-Bera (JB):                2.214
Skew:                           0.117   Prob(JB):                        0.331
Kurtosis:                       2.605   Cond. No.                         1.00
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Hedge Ratio is : 1.25

Portfolio 1 = ITUB4 - 1.25 * ITUB3

```
CADF Test Statistic for Portfolio 1  is  -2.97167715364
Critical Value is  {'1%': -3.4571053097263209, '5%': -2.873313676101283, '10%': -2.5730443824681606}
At 10% we  Reject H0 - Time series is stationary
At 5% we  Reject H0 - Time series is stationary
At 1% we  Accept H0 - Time series is not stationary
```

Regression Results for ITUB3 as dependent variable:

```
Regression Results for ITUB3 as dependent variable:

                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       1.000
Model:                            OLS   Adj. R-squared:                  1.000
Method:                 Least Squares   F-statistic:                 8.547e+05
Date:                Sun, 07 Jan 2018   Prob (F-statistic):               0.00
Time:                        19:02:16   Log-Likelihood:                 118.96
No. Observations:                 252   AIC:                            -235.9
Df Residuals:                     251   BIC:                            -232.4
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
x1             0.8011      0.001    924.512      0.000       0.799       0.803
==============================================================================
Omnibus:                        2.606   Durbin-Watson:                   0.679
Prob(Omnibus):                  0.272   Jarque-Bera (JB):                2.212
Skew:                          -0.117   Prob(JB):                        0.331
Kurtosis:                       2.605   Cond. No.                         1.00
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Hedge Ratio is : 0.8

Portfolio 2 = ITUB3 - 0.8* ITUB4

```
CADF Test Statistic for Portfolio 2  is  -2.97167715364
Critical Value is  {'1%': -3.4571053097263209, '5%': -2.873313676101283, '10%': -2.5730443824681606}
At 10% we  Reject H0 - Time series is stationary
At 5% we  Reject H0 - Time series is stationary
At 1% we  Accept H0 - Time series is not stationary
```

For the initial time range the selected portfolio with lower CADF t-stat is:

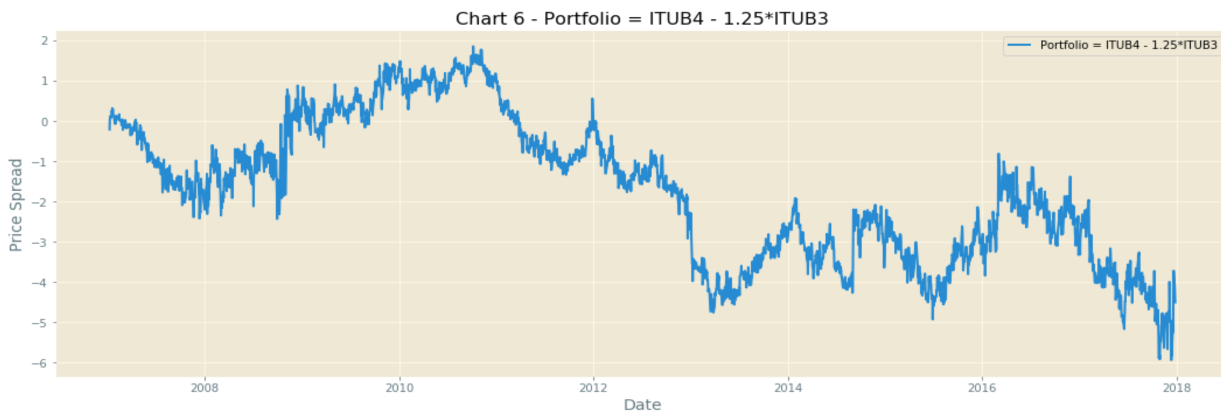
<u>Portfolio = ITUB4 - 1.25*ITUB3</u>

The CADF test results show that that we can Reject the null hypothesis that the portfolio is not stationary at a 5% level, since for both relationship the stat-t is the same we are going to arbitrarily select one of the portfolios, but the implemented test would return the lowest t-stat CADF result automatically.

For the first 252 days of data on the set the selected portfolio based on the lowest CADF test stats is :

Portfolio = ITUB4 - 1.25*ITUB3

The resulting historical spread for the remaining of the data range is demonstrated on Chart 6, visually it is possible to realize that the relationship which result on a cointegrated relationship based on the first 252 data points does not hold subsequently, nevertheless we will implement a trading strategy and analyses the results.



2.4 Implementation of the Strategy

2.4.1 Half-life of mean reversion

In *Algorithmic Trading: Winning Strategies and Their Rationale (Wiley Trading), Ernie Chan* suggests a Bollinger band approach where trades are entered when price deviates by more than x standard deviations from the mean, where x is a parameter to be optimized. The lookback period for the rolling mean and standard deviation can be optimized or set to the half-life of mean reversion.

Half-life can be estimated from a Ornstein-Uhlenbeck stochastic process, which is a mean-reverting process, described by the SDE:

$$dX_t = \alpha\,(\mu - X_t)\,dt + \sigma\,dW_t$$

where α > 0 and Wt is the Wiener process. It can easily be solved explicitly:

$$X_t = e^{-\alpha t} X_0 + \mu\,(1 - e^{-\alpha t}) + \int_0^t \sigma\, e^{\alpha(s-t)}\, dW_s.$$

So, we deduce that

$$\mathrm{E}X_t = e^{-\alpha t}X_0 + \mu\left(1 - e^{-\alpha t}\right) = \mu + o(1), \qquad \text{as } t \to \infty,$$
$$\operatorname{Var} X_t = \frac{\sigma^2}{2\alpha}\left(1 - e^{-2\alpha t}\right) = \frac{\sigma^2}{2\alpha}, \qquad \text{as } t \to \infty.$$

Half-life of the mean-reversion, $t_{1/2}$, is the average time it will take the process to get pulled half-way back to the mean. To this end, we consider the ODE x˙=α(μ−x)x˙=α(μ−x) , which has the solution x(t) = μ + e-αt (x0 – μ). So, we can find the half-time from the equation

x(t1/2)−μ=x0−μ2,x(t1/2)−μ=x0−μ2,

i.e. t1/2=log2αt1/2=log2α . The higher the mean-reversion speed is, the smaller is the half-life.

Calculating the half-life for the spread we got:

Half-Life for the spread-portfolio is: 74 Days

2.5 Trading Strategy

To implement an example trade strategy, we are going to use a technique known as Bollinger Bands*. It involves taking a rolling simple moving average (SMA) of a price series and then forming "bands" surrounding the series that are a scalar multiple of the rolling standard deviation of the price series. The lookback period for the moving average and standard deviation is identical and we are setting it to the calculated half-life above.

The trades will be generated based on calculated z-score (also known as a standard score) of the current latest spread price. This is achieved by taking the latest portfolio market price, subtracting the rolling mean and dividing by the rolling standard deviation. Once this z-score is calculated a position will be opened or closed out under the following conditions:

- `zscore < -zentry: Long entry`
- `zscore > +zentry: Short entry`
- `zscore >= -zexit: Long close`
- `zscore <= +zexit: Short close`

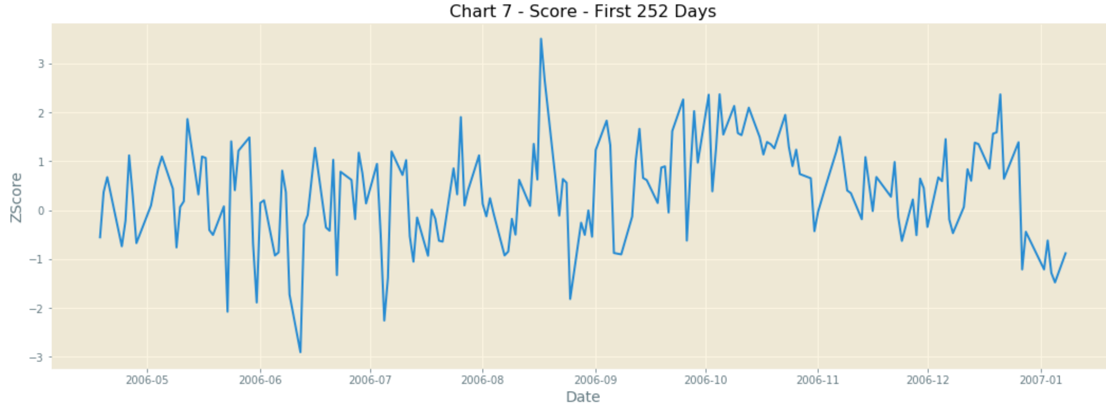
Where Zscore is the latest standardized spread price, Zentry is the trade entry threshold and Zexit is the trade exit threshold.

Note that if Zexit = 0, this means that a trade is closed when the price mean-reverts to the current mean (calculated as a rolling mean with lookback equal the half-life). If Zexit = Zentry, the position is closed when the prices move beyond the opposite band so as to as to trigger a trading signal of the opposite sign.

Since we are attributing the lookback period of the analysis to the **half-life** of the spread, to finish implementing the strategy we need to define **Zentry** and **Zexit**. For the initial example we are going to visually define it looking at the ZSpread chart for the first 252 days. It is possible to see that the Z_Spread varies most of the time between -2 and 2, so we are setting *Zentry=2* and *Zexit = -2*, so the close signal matches the enter signal, meaning we are always positioned on the spread.

Implementation described on Algorithmic Trading: Winning Strategies and Their Rationale (Wiley Trading) by Ernie Chan, and Advanced Algorithmic Trading by Michael L. Halls Moore

Chart 7 - Score - First 252 Days



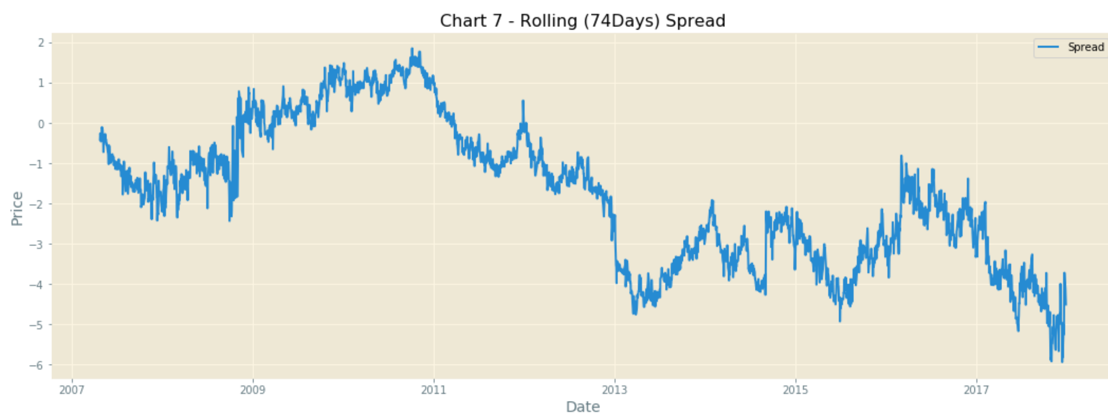
Generating the strategy based on the previous section explanation we got all entry and exit points for both Long and Short position, table below shows an extract of the strategy table per day:

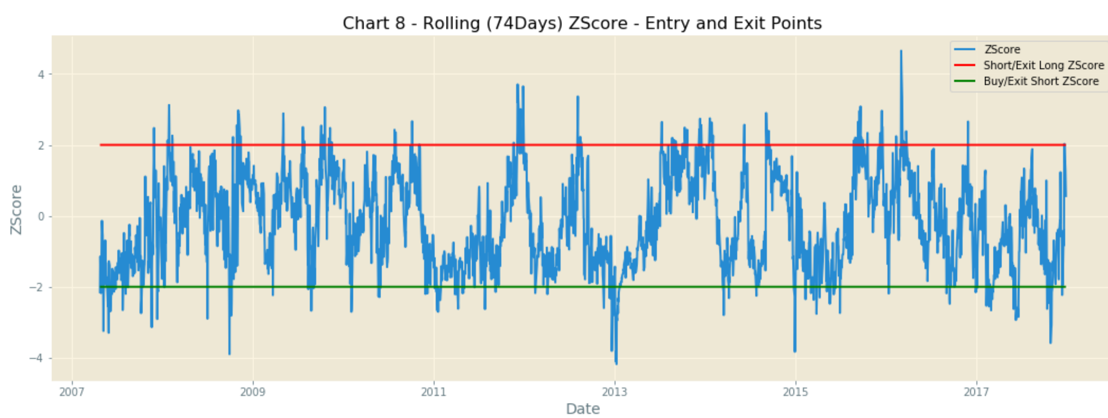
DATE	sprd	ITUB4	ITUB3	IBOV	z_score	long_entry	short_entry	long_exit	short_exit	positions_long	positions_short	positions
2007-04-26	-0.4400	13.46	11.12	49067.69	-2.179277	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2007-04-27	-0.2600	13.49	11.00	49229.60	-1.145132	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2007-04-30	-0.4600	13.19	10.92	48956.39	-2.179310	1.0	0.0	0.0	1.0	1.0	0.0	1.0
2007-05-02	-0.2925	13.57	11.09	49471.54	-1.216792	0.0	0.0	0.0	0.0	1.0	0.0	1.0
2007-05-03	-0.1625	13.90	11.25	50218.22	-0.483707	0.0	0.0	0.0	0.0	1.0	0.0	1.0
...	...	...	...	...	...	...	...	...	...	...	...	...
2017-12-21	-4.0175	42.72	37.39	75133.43	1.512672	0.0	0.0	0.0	0.0	1.0	0.0	1.0
2017-12-22	-3.7175	42.52	36.99	75186.53	2.028719	0.0	1.0	1.0	0.0	0.0	-1.0	-1.0
2017-12-26	-4.1100	42.69	37.44	75707.73	1.301140	0.0	0.0	0.0	0.0	0.0	-1.0	-1.0
2017-12-27	-4.3650	42.46	37.46	76072.54	0.838235	0.0	0.0	0.0	0.0	0.0	-1.0	-1.0
2017-12-28	-4.5300	42.57	37.68	76402.08	0.529596	0.0	0.0	0.0	0.0	0.0	-1.0	-1.0

2641 rows × 12 columns

```
Wall time: 517 ms
```

Charts 7 and 8 provides an easy visual way to understand how the strategy will develop for the rest of the data set, chart 7 is the historical spread chart (as seen before), chart 8 is the historical Z-Spread calculated using rolling standard deviations and mean using lookback period equal to the half-life (74 Days for the example). Green line indicates Open Long or Close Short and red line is the Open Short or Close Long points, so when Z-Spread touches one of the lines one of the 2 corresponding actions will happen.





2.5.1 Back Test of the Trading Strategy

Back test the strategy was a real challenge and it was hard to settle on proposed approaches that were found during research, some of the approaches proposed overly simplistic assumptions that produced really inexplicable results, and others highly complex that required a lot of data that it is not available easily, especially on an Emerging Market. The approach proposed is a NAV (Net Asset Value) based one with some simplifying assumptions but which i believe holds similarities to real execution:

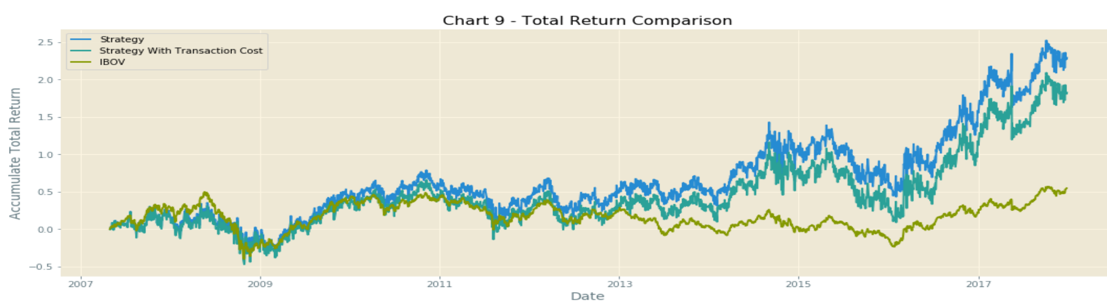
1. The trades are executed on the Close Price of the Assets
2. But, when an entry/exit point is flagged on the strategy, you can only execute the trade on the next close price, here the rational is that you used today's close price as an input on
Back test the strategy was a real challenge and it was hard to settle on proposed
3. The initial cash of the strategy is set to be equal to the unit value of the first trade of the strategy
4. There is no penalty for holding a "negative cash balance" but it discounts from the NAV Value
5. It is possible to trade fractions of the stocks.

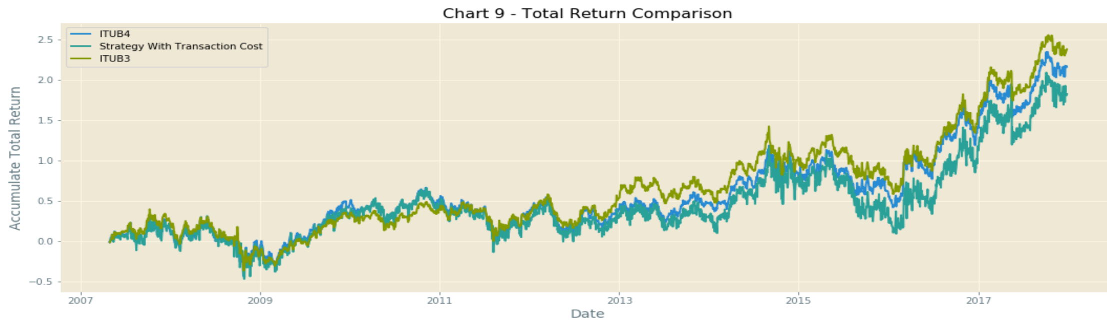
Following those rules, the daily NAV calculation is as follows:

1. Long Position: NAV is the value of your stock holdings
2. Short Position: NAV is zero since the cash proceeds from the sale balances the liabilities of the short holdings, but there is a daily adjustment to NAV based on asset/liability of the short position
3. Final NAV value daily: Final Cash of the Day + Long Position Value + Asset/Liability of the
4. Daily Strategy Return is (NAV/NAV(t−1))− 1

The interesting part of the approach is that given the Strategy Entry/Exit points defined on the previous step, the NAV calculation is almost entirely vectored which provides a very efficient calculation with transaction costs easily incorporated at each trade. On the example, 10 years of daily data were back tested in 1.13s

Chart 9 shows the return for the Strategy comparing the back test without transaction costs with the same back test with transaction cost of 0.5% per trade and with the Brazilian Benchmark Ibo Vespa Index.

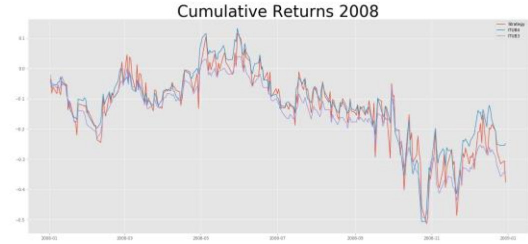




Slicing the returns per year and comparing to the assets (ITUB3 and ITUB4) and benchmark return IBOV the results are:

Cumulative Returns 2007

Cumulative Returns 2008



Cumulative Returns 2009

Cumulative Returns 2010

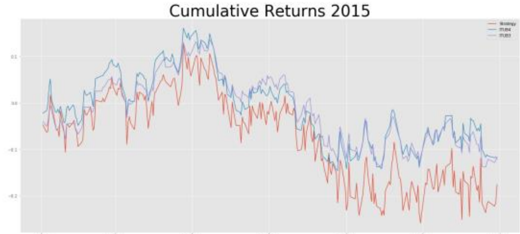
Cumulative Returns 2011

Cumulative Returns 2012

Cumulative Returns 2013

Cumulative Returns 2014

Cumulative Returns 2015

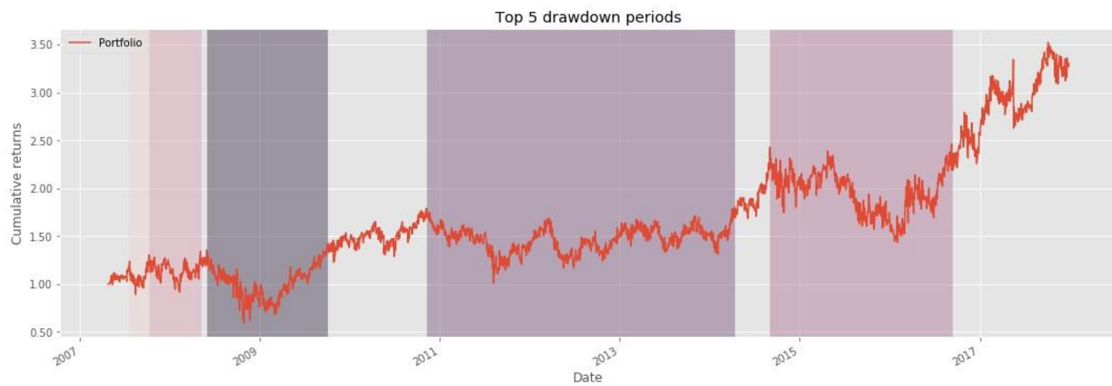


Cumulative Returns 2016

Cumulative Returns 2017

At this point it is possible to use a back test and risk/return analysis tool to study the behavior of the strategy. Applying some of the functionality of Pyplot (unfortunately not all functionality of Pyplot is working when you try to overwrite the standard benchmarks of the package), which makes some of the studies meaningless when benchmarked against S&P and other information non-related to the Brazil's market.

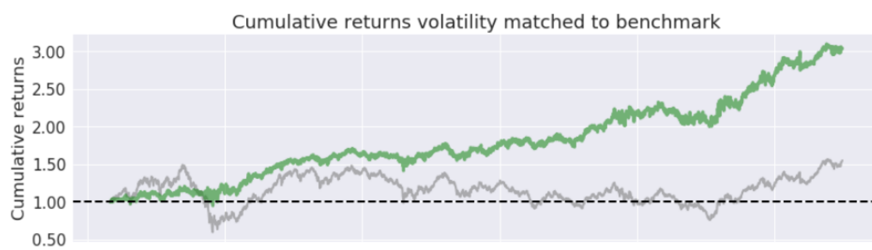
The Top 5 Draw Down periods:

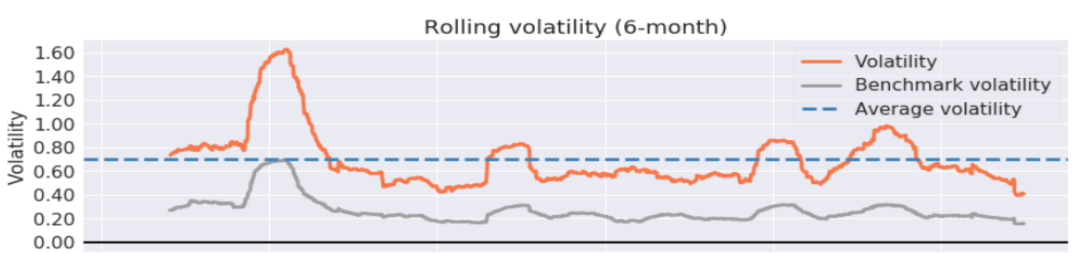


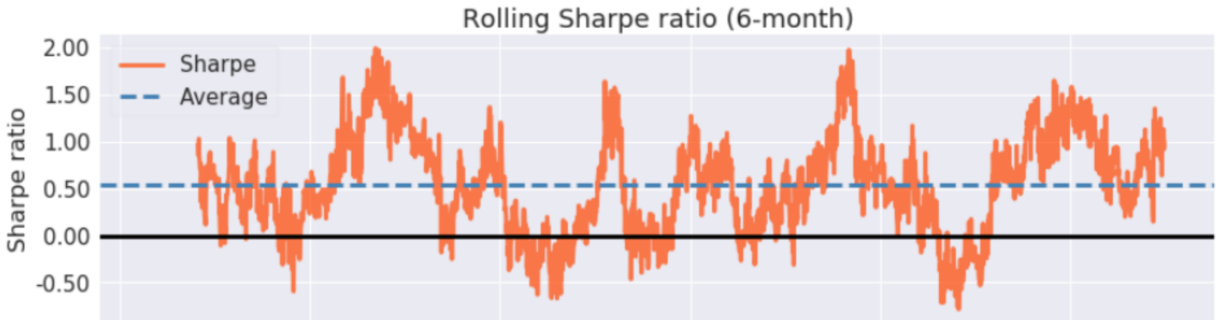
Some risk measures and risk adjusted returns from the pyplot package are demonstrate below:

	Backtest
Annual return	12.1%
Cumulative returns	230.0%
Annual volatility	72.7%
Sharpe ratio	0.52
Calmar ratio	0.21
Stability	0.76
Max drawdown	-56.4%
Omega ratio	1.10
Sortino ratio	0.78
Skew	0.45
Kurtosis	5.31
Tail ratio	1.08
Daily value at risk	-9.0%
Alpha	0.23
Beta	1.82

Worst drawdown periods	Net drawdown in %	Peak date	Valley date	Recovery date	Duration
0	56.41	2008-05-30	2008-10-28	2009-09-30	349
1	43.67	2010-11-08	2011-08-08	2014-04-11	895
2	41.03	2014-09-03	2016-02-02	2016-09-08	527
3	29.73	2007-10-09	2008-02-11	2008-05-05	150
4	27.99	2007-07-23	2007-08-16	2007-10-09	57







2.5.2 Conclusion for the initial set up

The returns for the first proposed approach were able to track the main Brazilian benchmark until 2014 when the strategy started to outperform it even when transactions cost is considered. Comparing to buy and hold strategy of the main components the strategy as able to keep pace with them until 2010 when it started to lag. Although the performance was good especially on the last 3 years overall the volatility was big was well producing a Sharpe Ratio of 0.52. We saw the biggest draw down during the 2008 financial crisis were the strategy lost 56.41% for 349 days.

2.6 Optimizing the Strategy

2.6.1 Look Ahead Bias

One important point to avoid on the set up of the strategy is the **_Look Ahead Bias_** which is the bias created using information or data in a study or simulation that would not have been known or available during the period being analyzed. This will usually lead to inaccurate results in the study or simulation. Look ahead bias was introduced on the last strategy on purpose, the half-life calculation utilized the information from the entire dataset, which would not be available at the time of implementation.

The objective of the following sections is to propose an optimized strategy that avoids any **look ahead bias**.

2.6.2 Proposed Optimized Approach

Look Back ranges for the Engle-Granger implementation:

Since for the entire dataset the T-Stat results for the CADF test produces similar results using ITUB4 or ITUB3 as dependent variable the strategy will use ITUB4 as dependent and ITUB3 as independent variables for the entire strategy. The strategy them will follow the procedure:

1. Rebalance the portfolio every 30 days

2. The hedge ratio will be set initial 90 days of data and will be fixed for the rest of the strategy (the initial objective was to optimize the hedge ratio according to the latest test, but although it is easy to generate the strategy it proved a challenge to incorporate the rebalances on the back-test algorithm.

3. The half-life will be calculated using the initial 90 days of data prior to the rebalance date (during the research for this chapter it comes to the attention the short ranges of data used to estimate the half-life produces very short or very long Half Life, so we are setting a limiting interval of 20 to 100 days

4. If CADF results produces a cointegrated relationship using the last 90 days of data, the strategy until the next rebalance will be set to Zentry=1Zentry=1 and Zexit=−1Zexit=−1 (as described earlier meaning that we are always positioned on the spread)

5. If CADF results produces a non-cointegrated relationship using the last 90 days of data, the strategy until the next rebalance will be set to Zentry=2Zentry=2 and Zexit=0Zexit=0 (meaning that we are more conservative on the entry points and we are exiting earlier the positions)

The initial adjustments according to the first 90 days of data are :

```
CADF T-Stat =  -2.74097799238 Suggesting Cointegration for the selected Range
Selected Portfolio is =  Portfolio = ITUB4 - 1.24*ITUB3
Calculated Half Life is :  1  Adjusted Half Life is : 20
Initial Z_Entry :  1 Initial Z_Exit :  : -1
```

With the initial parameters defined we can run the strategy for the rest of the data set, bellow is a extract of the strategy daily flags.

DATE	ITUB4	ITUB3	IBOV	sprd	z_score	long_entry	short_entry	long_exit	short_exit	positions_long	positions_short	positions
2006-05-15	11.01	8.87	39271.45	0.0112	0.078662	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2006-05-16	11.12	8.87	39416.44	0.1212	1.025733	0.0	1.0	1.0	0.0	0.0	-1.0	-1.0
2006-05-17	10.93	8.72	38290.68	0.1172	0.900024	0.0	0.0	0.0	0.0	0.0	-1.0	-1.0
2006-05-18	10.73	8.72	37807.15	-0.0828	-1.080269	1.0	0.0	0.0	1.0	1.0	0.0	1.0
2006-05-19	10.52	8.56	37732.86	-0.0944	-1.088126	1.0	0.0	0.0	1.0	1.0	0.0	1.0
...	...	...	...	...	...	...	...	...	...	...	...	...
2017-08-16	39.60	34.83	68594.30	-3.5892	-0.396930	0.0	0.0	0.0	0.0	0.0	-1.0	-1.0
2017-08-17	39.28	34.50	67976.80	-3.5000	-0.130698	0.0	0.0	0.0	0.0	0.0	-1.0	-1.0
2017-08-18	39.66	35.09	68714.66	-3.8516	-1.508027	1.0	0.0	0.0	1.0	1.0	0.0	1.0
2017-08-21	39.37	34.64	68634.65	-3.5836	-0.476420	0.0	0.0	0.0	0.0	1.0	0.0	1.0
2017-08-22	40.21	35.37	70011.25	-3.6488	-0.715165	0.0	0.0	0.0	0.0	1.0	0.0	1.0

2790 rows × 12 columns

Now we can just run it on our back-test algorithm to see if we got improved returns on the proposed optimizations, we set transaction cost to 0.5% of the financial value on every trade.

An extract of the resulting back test is shown next:

exit	positions_long	...	entry_long_cash	entry_short_cash	exit_long_cash	exit_short_cash	initial_cash	end_cash	long_nav	short_nav	final_cash	nav
0.0	0.0	...	0.00000	0.0	0.0	0.00000	10.580000	0.0	0.0000	0.00	10.580000	10.580000
0.0	0.0	...	0.00000	0.0	0.0	0.00000	10.580000	0.0	0.0000	0.00	10.580000	10.580000
0.0	0.0	...	0.00000	0.0	0.0	0.00000	10.580000	0.0	0.0000	0.00	10.580000	10.580000
1.0	1.0	...	0.00000	0.0	0.0	0.00000	10.580000	0.0	0.0000	0.00	10.580000	10.580000
1.0	1.0	...	-10.57260	0.0	0.0	0.00000	10.580000	0.0	10.5200	0.00	0.007400	10.527400
...	...	...	...	...	...	...	...	...	...	...	...	...
0.0	0.0	...	0.00000	0.0	0.0	0.00000	-27.466262	0.0	43.5375	0.32	-27.466262	16.391238
0.0	0.0	...	0.00000	0.0	0.0	0.00000	-27.466262	0.0	43.1250	-0.38	-27.466262	15.278738
1.0	1.0	...	0.00000	0.0	0.0	0.00000	-27.466262	0.0	43.8625	0.29	-27.466262	16.686238
0.0	1.0	...	-39.56685	0.0	0.0	43.37205	-27.466262	0.0	39.3700	0.00	-23.661062	15.708938
0.0	1.0	...	0.00000	0.0	0.0	0.00000	-23.661062	0.0	40.2100	0.00	-23.661062	16.548938

Comparing the accumulated returns between the 2 strategies it is possible to see that the new one improved the results especially after 2011, visually it seems that the volatility profile is the same which can be analyzed by comparing results from the pyplot package

Chart 9 compares the 2 accumulated returns

2.7 Final Conclusions

In this chapter we simulated a simple long/short strategy which is based on a cointegrated series, the series itself is the product of a linear combination between 2 assets. Different share classes from the same company are suitable for this type of analysis, especially in emerging markets which this type of situation tends to take more time to be adjusted, in our example we simulated 10 years of the strategy with very reasonable results. Of course, in practice a lot of the assumptions don't hold and the performance can be very different from the one predicted by the model, but still it is a relatively simple strategy to be implemented which i think compensates the time spent on researching this kind of opportunities.

A further development of this strategy would be the addition of a third asset which we believe holds a relationship with the basic strategy, for our example it is accepted by the market that the difference between the ON and PN share classes in Brazil are due to liquidity differences in the assets (as explained on the introduction) a natural step would be the introduction of a asset (preference for a tradable one) that could be used as a proxy of market liquidity.

For the pair of shares used in the analysis we used the CADF framework for stationarity test, one of the biggest drawbacks of the test was that it was only capable of being applied to two separate time series. With the addition of a third asset we need another procedure for stationarity test, for it we would use Johansen that allows us to determine if three or more-time series are cointegrated. We will then form a stationary series by taking a linear combination of the underlying series and them create the strategy.

3.1 References

CVA Calculation for an Interest Rate Swap

- *Options, Futures, and Other Derivatives2011 by John C. Hull*
- *Paul Wilmott on Quantitative Finance 2nd Edition Hardcover by Paul Wilmott*
- *Counterparty Credit Risk and Credit Value Adjustment: A Continuing Challenge for Global Financial Markets,by Jon Gregory*

Cointegration and Arbitrage Trading on Dual-Class Shares

- *Algorithmic Trading: Winning Strategies and Their Rationale by Ernie Chan*
- *Advanced Algorithmic Trading by Michael L. Halls Moore*