Master in Financial Engineering
Financial Econometrics

Case study 1: Modeling a financial time series, cointegration, and pair trading (Statarb)

Pr. Florian Pelgrin - T.A. Maxime Couvert

Ecole Polytechnique Federale de Lausanne

Feb. 2019 - June 2019

Case study 1

- Exercise 1 : Modeling the 5-year interest rate
- Exercise 2 : The term structure of interest rates
- Exercise 3 : Pair trading or Statarb

Exercise 2 : The term structure of interest rates

Road map of exercise 2

1. Exercise 2
   - Part I : OLS estimation
   - Part II : Error correction model

2. Exercise 3 : Pair trading or Statarb

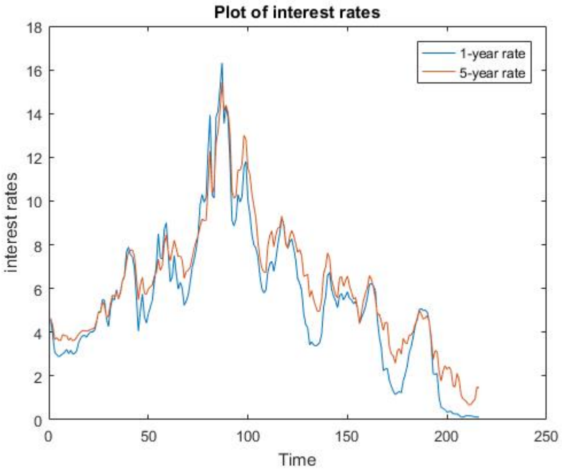
Aim of the exercise

- Introduce the concept of cointegration
- Use cointegration to implement an error correction model (ECM)

Part I

- X_t : 1-yr and 5-year interest rate series
- Period : 1960Q1-2013Q3
- Goal : assess whether the two series are cointergrated
- Three steps :
  1. Identify whether the two series are stationary or not ;
  2. Identify a linear relation between the two series ;
  3. Test whether this linear relation is stationary or not.

Plot the time series Xt



OLS estimation of the term structure

The expectations hypothesis of the term structure of interest rates claims that the following relationship between the m-year and n-year interest rate is satisfied (with $m < n$) :

$$y_t^{(n)} = \alpha + \beta y_t^{(m)} + u_t \tag{1}$$

where u_t is the (i.i.d.) error term, and α is the term premium.
$\rightarrow$ The pure expectations hypothesis predicts that $\beta = 1$.

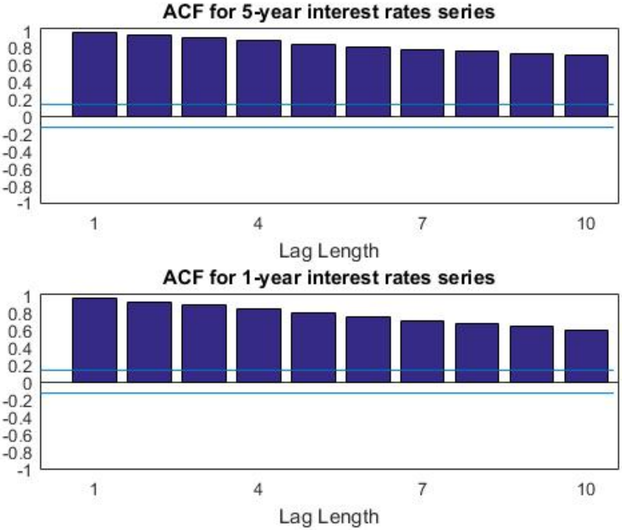
OLS estimation of the term structure

Figure – OLS Estimation

	Estimate	SE	tStat	pValue
(Intercept)	1.3059	0.092981	14.045	3.456e-32
x1	0.88452	0.014626	60.476	1.5929e-136

```
Number of observations: 216, Error degrees of freedom: 214
Root Mean Squared Error: 0.693
R-squared: 0.945,  Adjusted R-Squared 0.944
F-statistic vs. constant model: 3.66e+03, p-value = 1.59e-136
```

ACF of the dependent and explanatory variables



Reminder on Augmented Dickey-Fuller test (ADF)

- **Goal :** Assess whether there exists a root on the unit circle in the autoregressive lag polynomial of an AR(p) specification.
- **Design :** The ADF test is based on estimating the following test regression :

$$X_t = \sum_{j=1}^{p} \rho_j X_{t-j} + z_t' \delta + \epsilon_t$$

where (ϵ_t) is a weak white noise, z_t is a set of exogenous regressors ($z_t = \{1, t\}, z_t = \{t\}$, or $z_t = \{1\}$), p is the lag order (that can be determined by differen techniques) and $(\rho_1, \cdots, \rho_p)'$ and δ are parameters to estimate.

Reminder on Augmented Dickey-Fuller test (ADF)

- **Design :** Using the so-called Beveridge-Nelson decomposition, the test regression can be rewritten as follows

$$X_t = (1 - \Phi(1))\, X_{t-1} + \sum_{j=1}^{p-1} \alpha_j \Delta X_{t-j} + z_t' \delta + \epsilon_t$$

where $\Phi(1) = 1 - \sum_{j=1}^{p} \phi_j$.

- **Test specification :**

$$H_0 : \text{The series is non-stationary} \rightarrow \Phi(1) = 1$$
$$H_a : \text{The series is "weakly stationary"} \rightarrow \mid \Phi(1) \mid < 1.$$

Augmented Dickey-Fuller test (ADF)

- **Assumptions :**
  - $z_t = 1$ for all t
  - p=3
- **Test :**

TABLE – ADF test

	5-year rate	1-year rate
p-values	0.4758	0.3191

Quick reminder of the theory - cointegration

Definition

If X_{1t} and X_{2t} are integrated of order 1 but there exists a linear combination

$$Z_t = \beta_1 X_{1t} + \beta_2 X_{2t}$$

such that Z_t is integrated of order zero, then X_{1t} and X_{2t} are said to be cointegrated.

Cointegration

- Is the 5-year interest rate non-stationary ?
- Is the 1-year interest rate non-stationary ?
- Are the residuals stationary ?

TABLE – ADF test on residuals

	residuals
p-value	0.0018

Reminder on the error correction model

Two-step method of Engle and Granger :

- A simple two-step procedure can be implemented as follows :
  1. Estimate Eq. 1 by OLS ;
  2. Compute the OLS residuals and especially $\hat{u}_{t-1}$;
  3. Conduct the following OLS regression :

$$\Delta y_t^{(5y)} = c + \lambda \hat{u}_{t-1} + \sum_{j=1}^{k_1} a_j \Delta y_{t-j}^{(5y)} + \sum_{j=1}^{k_2} b_j \Delta y_{t-j}^{(1y)} + \epsilon_t$$

where the parameters are c, λ (¡ 0), $\{a_j\}_{j=1,\cdots,k_1}$, and $\{b_j\}_{j=1,\cdots,k_2}$, and the lag orders k_1 and k_2 can be determined by using information criteria or a general-to-specific procedure (with $k_1 \leq 3$ and $k_2 \leq 3$).

Choice of the error correction model

Two-step method of Engle and Granger : Choose (k_1, k_2) thanks to AIC and BIC :

TABLE – AIC and BIC

Model	AIC	BIC
(1,1)	359.3693	372.8332
(2,1)	351.1588	367.9653
(3,1)	346.0870	366.2265
(1,2)	350.4173	367.2238
(2,2)	352.1987	372.3664
(3,2)	347.2370	370.7331
(1,3)	346.1629	366.3024
(2,3)	347.8074	371.3035
(3,3)	349.0870	375.9397

Estimation of Error Correction Model

FIGURE – ECM :$(k_1 = 3,\ k_2 = 1)$

	Estimate	SE	tStat	pValue
(Intercept)	-0.0081771	0.037105	-0.22038	0.82579
x1	-0.12253	0.057509	-2.1307	0.034304
x2	0.31841	0.14258	2.2332	0.026614
x3	-0.05513	0.11091	-0.49709	0.61966
x4	-0.2338	0.069338	-3.3718	0.0008917
x5	0.17391	0.069045	2.5188	0.012537

Number of observations: 212, Error degrees of freedom: 206
Root Mean Squared Error: 0.54
R-squared: 0.128, Adjusted R-Squared 0.107
F-statistic vs. constant model: 6.05, p-value = 3.02e-05

Exercise 3 : Pair trading or Statarb

Exercise 3 : Pair trading or Statarb

Consider a pair of daily stock prices. Suppose the two price are cointegrated.

- How can we use an ECM to implement a "simple" financial strategy ?
- Find such a pair among your favorite stocks in a given industry, and estimate the corresponding model.