Presentation about Statistical Arbitrage (Stat-Arb), using Cointegration on the Equity Market

By Yoann BOURGEOIS and Marc MINKO

Derivative Models Review Group (DMRG-Paris)
HSBC-CCF

DMRG-Paris - 17/06/05

CCF

PLAN

- ❑Introduction
- ❑Part I: Mathematical Framework
- ❑Part II: Description of the proposed strategy and first results
- ❑Conclusion

Introduction

- Single stocks in the Equity Market generally are not stationary.
- But, their yields, in many cases are.
- From the econometrical point of view, they are generally told to be Integrated of order 1.
- Cointegration is a mathematical theory that helps to handle the problem generated by non-stationary data.
- With the help of this theory, we propose to build linear combinations of these single stocks that are stationary.
- Such combinations can be traded and are called synthetic assets.
- Eventually, these stationary assets have the mean reversion property and we will use this property in order to set up arbitrage strategies.

- ❑Description of the framework of our strategy
- ❑Statistical Analysis of models that are Integrated of order 1 (ie I(1))

Description of the framework of our strategy (1)

- **Reminder about Vector AutoRegressive models (VAR)**

In what follows, we consider a VAR process X_t (p×1), which can be written:

$$X_t = \sum_{i=1}^{k} \Pi_i X_{t-i} + \Phi D_t + \varepsilon_t$$

with $1 \leq t \leq T$

$$D_t = (1, t, t^2)'$$

$X_{-k+1}, \ldots, X_0$ known

ε_t i.i.d with law $N_p(0, \Omega)$

Remark: **here, we suppose that the errors are i.i.d with a gaussian law, but it can easily be generalised with errors i.i.d with finite moments of order 2.**

Description of the framework of our strategy (2)

Definition: **Let's introduce the characteristic polynomial: $|A(z)|$ with**

$$A(z) = I - \sum_{i=1}^{k} z^i \Pi_i$$

Inversibility THEOREM for a VAR process

The VAR process X_t can be written as a function of its initial values and of the errors:

$$X_t = \sum_{s=1}^{k} C_{t-s}(\Pi_s X_0 + \ldots + \Pi_k X_{-k+s}) + \sum_{j=0}^{t-1} C_j(\varepsilon_{t-j} + \Phi D_{t-j})$$

with $C_0 = I$ et $\forall n \geq 1:\ C_n = \sum_{j=1}^{k \wedge n} C_{n-j}\Pi_j$. Let $C(z) \stackrel{\Delta}{=} \sum_{n \geq 0} z^n C_n$.

Then, $\exists \delta > 0 /$: this serie converges and inside the disc of radius δ:

$C(z)A(z) = I$ ie $C(z) = A(z)^{-1}$

Description of the framework of our strategy (3)

<u>Remark:</u> **the solution given by this theorem is valid whatever the parameters are. On the contrary, it is reminded in what follows that the parameters have to be constrained in order to define a stationary VAR process.**

<u>Definition:</u> **a process** X_t **is told strongly stationary iif**

$$\forall h \geq 1: \quad \mathrm{Law}(X_{t_1},\dots,X_{t_m}) = \mathrm{Law}(X_{t_1+h},\dots,X_{t_m+h})$$

It is told weakly stationary of order 2 iif

$$EX_t = \text{constant} \;\text{ et } \mathrm{Var}X_t = \text{constant}$$

<u>Remark:</u> **in part I, strong stationarity is used since the errors are gaussian.**

Description of the framework of our strategy (4)

FUNDAMENTAL HYPOTHESIS :

$$|A(z)| = 0 \Rightarrow |z| > 1 \text{ or } z = 1$$

Remark: this fundamental hypothesis excludes explosive roots with |z|<1 as well as seasonal roots (|z|=1 and z different from 1). If z=1 is a root, then the process is told to have a unit root.

Description of the framework of our strategy (5)

THEOREM defining the necessary and sufficient condition for the stationarity of a VAR process

Under the fundamental hypothesis, a necessary and sufficient condition for $X_t - EX_t$ to be stationary is $|A(1)| \neq 0$. In such a case, the MA representation of the VAR process is obtained :

$$X_t = \sum_{n \geq 0} (C_n \varepsilon_{t-n} + \Phi D_{t-n}) \quad \text{where } \exists \delta > 0 / : \; C(z) = \sum_{n \geq 0} z^n C_n = A(z)^{-1}$$

is convergent for $|z| < 1 + \delta$

Remark: **(i) when $\Phi = 0$, one recognizes the WOLD theorem.**

(ii) it is checked that with gaussian errors, the strong stationarity is recovered, whereas in the general case, the weak stationarity is obtained.

Description of the framework of our strategy (6)

- **Basic definitions for cointegration**

Preliminary remark:

Many economic variables are non stationary and the kind of non-stationarity that is considered here can be removed by one or several differentiations. In what follows, we will suppose that:

$$\varepsilon_t \text{ is i.i.d with law } \mathrm{N_p}(0,\Omega)$$

Definition:

$$\text{a process } \mathrm{Y_t} / : \mathrm{Y_t} - EY_t = \sum_{i \geq 0} C_i \varepsilon_{t-i} \text{ is integrated}$$

$$\text{of order } 0 \overset{\Delta}{\Leftrightarrow} C = \sum_{i \geq 0} C_i \neq 0$$

Description of the framework of our strategy (7)

Remarks:

(i) C may be singular and in fact this is a pathway to cointegration.

(ii) The process defined in the stationarity theorem is I(0) and in fact $C(1) = A(1)^{-1}$ is regular.

(iii) In dimension 1, stationarity and I(0) processes define the same concepts.

Definition:

a process X_t is told "integrated of order d"

(and noted down I(d), $d \geq 1$) $\overset{\Delta}{\Leftrightarrow} \Delta^d (X_t - EX_t)$ is I(0)

Description of the framework of our strategy (8)

Remark: the property of being integrated is connected with the stochastic part of the process since the mean is substracted from the process in the definition. The concept of I(0) process is defined without considering deterministic terms such as the mean or the trend.

Définition:

Let X_t : I(1). X_t is told cointegrated with the cointegration vector $\beta \neq 0 \overset{\Delta}{\Leftrightarrow} \beta' X_t$ is stationary.

The cointegration rank is the number of cointegration relations that are linearily independant.

Last, the vectorial space that is generated by cointegration relations is the cointegration space.

Description of the framework of our strategy (9)

❑ The Vector Error Correction Model (VECM) representation

We write the already used VAR model in a new way that is the VECM, because this is the model that is used in the cointegration theory.

Every VAR(k) can be written : $\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Phi D_t + \varepsilon_t$

with $\Pi = \sum_{i=1}^{k} \Pi_i - I$ and $\forall i \geq 1 :\ \Gamma_i = -\sum_{j=i+1}^{k} \Pi_j$

The following quantity is also used : $\Gamma \overset{\Delta}{=} I - \sum_{i=1}^{k-1} \Gamma_i$

The characteristic polynomial of X_t can be written : $A(z) = (1-z)I - z\Pi - \sum_{i=1}^{k-1} (1-z) z^i \Gamma_i$

Let's note that : $A(1) = -\Pi$ and $\dot{A}(1) = \frac{d}{dz} A(z)\Big|_{z=1} = -\Pi - I + \sum_{i=1}^{k-1} \Gamma_i = -\Pi - \Gamma$

Description of the framework of our strategy (10)

GRANGER Representation THEOREM (1)

If $|A(z)| = 0 \Rightarrow |z| > 1$ or $z = 1$ and if $\text{rank}(\Pi) = r < p$

then $\exists \alpha, \beta$ $(p \times r)$ with rank r so that : $\Pi = \alpha\beta^{'}$

A necessary and sufficient condition for $\Delta X_t - E\Delta X_t$ and

$\beta^{'} X_t - E\beta^{'} X_t$ to be stationary is that $\left| -\alpha^{'}_{ortho} \dot{A}(1)\beta_{ortho} \right| = \left|\alpha^{'}_{ortho}\Gamma\beta_{ortho}\right| \neq 0$

In this case, X_t can be written with the MA representation :

$$X_t = C\sum_{i=1}^{t}(\varepsilon_i + \Phi D_i) + C_1(L)(\varepsilon_t + \Phi D_t) + A$$

where A depends on the initial values and is so that : $\beta^{'}A = 0$

$$C = \beta_{ortho}(\alpha^{'}_{ortho}\Gamma\beta_{ortho})^{-1}\alpha^{'}_{ortho}$$

X_t is clearly an $I(1)$ process that is cointegrated with the r column vectors of β

Description of the framework of our strategy (11)

GRANGER Representation THEOREM (2)

Last, the serie $C_1(z)$ is so that : $A^{-1}(z) = C\dfrac{1}{1-z} + C_1(z)$, with $z \neq 1$ and with $1+\delta$ as convergence radius $(\delta > 0)$

Remark: **clearly, from the MA writing,** $\beta' X_t - E\beta X_t$ **is stationary since** $\beta' C = 0$ et $\beta' A = 0$ **. Besides,** $\beta' C_1(L)(\varepsilon_t + \Phi D_t)$ **is a representation of the distance of** $\beta' X_t - E\beta' X_t$ **from the balance position. The relation** $\beta' X_t = E\beta' X_t$ **defines underlying economic relations and supposes that all agents react to the distance from the balance position through the adjustment coefficient** α **and make the variables satisfy the economic relations again.**

<u>Remark (end):</u> **it has to be noted that the relations** $\beta^{'} X_t = E\beta^{'} X_t$ **are not asymptotic balance relations with t→+∞ or else relations between the levels of variables in balance. It should be told instead that these relations are relations between the portfolio variables that are described by the statistical model and that translate the adjustment behaviour of the agents.**

Statistical Analysis of I(1) models - (1)

- **The existence of the cointegration vectors, which is also known as the Reduced Rank hypothesis, is expressed in a parametric form, so that the Likelihood method can be applied.**
- **Therefore, estimators and statistical tests related to a fixed number of cointegration vectors can be written with closed formula.**

- Let's consider the following general VECM model:

$$\Delta X_t = \alpha\beta' X_{t-1} + \sum_{i=1}^{k} \Gamma_i \Delta X_{t-i} + \Phi D_t + \varepsilon_t$$

with $1 \leq t \leq T$,

ε_t i.i.d with law $N_p(0,\Omega)$

and $(\alpha, \beta, \Gamma_1, \ldots, \Gamma_k, \Phi, \Omega)$ as free parameters

CCF

Statistical Analysis of I(1) models - (3)

- **As already told, an analysis of the likelihood function is done with the following notation:**

$$Z_{0t} = \alpha\beta' Z_{1t} + \Psi Z_{2t} + \varepsilon_t$$

with

$$\begin{aligned} Z_{0t} &= \Delta X_t \\ Z_{1t} &= X_{t-1} \\ Z_{2t} &= (\Delta X_{t-1}^{'}, \ldots, \Delta X_{t-k}^{'}, D_t^{'})^{'} \\ \Psi &= (\Gamma_1, \ldots, \Gamma_k, \Phi) \end{aligned}$$

Statistical Analysis of I(1) models - (4)

Let's introduce:

$$M_{ij} = \frac{1}{T}\sum_{t=1}^{T} Z_{it} Z'_{jt}$$

$$\text{with} \quad 0 \leq i, j \leq 2$$

<u>Remark:</u>

$$M_{ij} = M'_{ji}$$

With a constant, the log-likelihood can be written:

$$\ln L(\Psi, \alpha, \beta, \Omega) = -\frac{T}{2}\ln|\Omega|$$

$$-\frac{1}{2}\sum_{t=1}^{T}(Z_{0t} - \alpha\beta' Z_{1t} - \Psi Z_{2t})'\Omega^{-1}(Z_{0t} - \alpha\beta' Z_{1t} - \Psi Z_{2t})$$

Statistical Analysis of I(1) models - (5)

- **First order conditions give for Ψ :**

$$\sum_{t=1}^{T}(Z_{0t}-\alpha\beta' Z_{1t}-\hat{\Psi} Z_{2t})Z_{2t}' = 0$$

$$\hat{\Psi}(\alpha,\beta) = M_{02}M_{22}^{-1} - \alpha\beta' M_{12}M_{22}^{-1}$$

- **The residuals R_{0t} and R_{1t} are defined by:**

$$R_{0t} = Z_{0t} - M_{02}M_{22}^{-1}Z_{2t}$$

$$R_{1t} = Z_{1t} - M_{12}M_{22}^{-1}Z_{2t}$$

(these residuals would be obtained while regressing respectively ΔX_t and X_{t-1} on

$$\Delta X_{t-1},\ldots,\Delta X_{t-k}$$

)

Statistical Analysis of I(1) models - (6)

Therefore the log-likelihood can be written:

$$\ln L(\alpha,\beta,\Omega) = -\frac{T}{2}\ln|\Omega| - \frac{1}{2}\sum_{t=1}^{T}(R_{0t}-\alpha\beta' R_{1t})'\Omega^{-1}(R_{0t}-\alpha\beta' R_{1t})$$

Let:

$$S_{ij} = \frac{1}{T}\sum_{t=1}^{T} R_{it}R_{jt}' = M_{ij} - M_{i2}M_{22}^{-1}M_{2j}$$

$$\text{with} \quad 0 \leq i,j \leq 1$$

For β **fixed, it is easy to infer** α **and** Ω **while regressing** R_{0t} **on** $\beta' R_{1t}$

So:

$$\hat{\alpha}(\beta) = S_{01}\beta(\beta' S_{11}\beta)^{-1}$$

$$\hat{\Omega}(\beta) = S_{00} - S_{01}\beta(\beta' S_{11}\beta)^{-1}\beta' S_{10} = S_{00} - \hat{\alpha}(\beta)(\beta' S_{11}\beta)\hat{\alpha}(\beta)'$$

Statistical Analysis of I(1) models - (7)

Therefore:

$$L_{\max}^{-\frac{2}{T}}(\beta) = \left|\hat{\Omega}(\beta)\right| = \frac{|S_{00}|}{|\beta' S_{11}\beta|}\left|\beta'(S_{11} - S_{10}S_{00}^{-1}S_{01})\beta\right|$$

and the FUNDAMENTAL THEOREM of the STATISTICAL ANALYSIS of I(1) models can be deduced:

Under hypothesis : $H(r) : \Pi = \alpha\beta'$, the MLE of β is given while solving the following equation : $\left|\lambda S_{11} - S_{10}S_{00}^{-1}S_{01}\right| = 0$

with the eigenvalues : $1 > \hat{\lambda}_1 > \ldots > \hat{\lambda}_p > 0$ and the

eigenvectors : $\hat{V} = (\hat{v}_1, \ldots, \hat{v}_p)$ normalized by $\hat{V}' S_{11} \hat{V} = I$

Statistical Analysis of I(1) models - (8)

Cointegration relations are infered by : $\hat{\beta} = (\hat{v}_1, \ldots, \hat{v}_r)$

and the maximized likelihood function can be written :

$L_{\max}^{-\frac{2}{T}}(H(r)) = |S_{00}| \prod_{i=1}^{r} (1 - \hat{\lambda}_i)$. The estimators of

the other parameters are obtained while inserting

$\hat{\beta}$ in the equations above, ie with $\beta = \hat{\beta}$ in the OLS.

The likelihood test : H(r) : $\text{rank}(\beta) = \text{r}$ against

$\text{rank}(\beta) > \text{r}$ has for statistic : $-\text{T} \sum_{\text{i=r+1}}^{\text{p}} \ln(1 - \hat{\lambda}_{\text{i}})$

Statistical Analysis of I(1) models - (9)

Remarks: (i) The r biggest eigenvalues are useful for getting the cointegration relations, while the p - r smallest are used in the JOHANSEN statistical test.

(ii) $\hat{\alpha} = S_{01}\hat{\beta}$

$$\hat{\beta}_{ortho} = S_{11}(\hat{v}_{r+1},\ldots,\hat{v}_{p})$$

$$\hat{\alpha}_{ortho} = S_{00}^{-1}S_{01}(\hat{v}_{r+1},\ldots,\hat{v}_{p})$$

Statistical Analysis of I(1) models - (10)

- **Models with constrained determinist terms**

  Up to now, the coefficients of Φ were totally free.

  From now, we shall also consider the case when the dominant coefficient is constrained. Therefore, we get two other models:

  with constrained constant:

$$\Delta X_t = \alpha(\beta' X_{t-1} + \rho) + \sum_{i=0}^{k} \Gamma_i \Delta X_{t-i} + \varepsilon_t$$

  with constrained linear trend:

$$\Delta X_t = \alpha(\beta' X_{t-1} + \rho t) + \sum_{i=0}^{k} \Gamma_i \Delta X_{t-i} + \mu + \varepsilon_t$$

Statistical Analysis of I(1) models - (11)

The same likelihood method can be used with the two new models. Only the notation differ.

With a constrained constant:

$$Z_{0t} \text{ becomes } Z^{*}_{0t} = \Delta X_t$$

$$Z_{1t} \text{ becomes } Z^{*}_{1t} = (X^{'}_{t-1}, 1)^{'}$$

$$Z_{2t} \text{ becomes } Z^{*}_{2t} = (\Delta X^{'}_{t-1}, \ldots, \Delta X^{'}_{t-k})^{'}$$

With a constrained linear trend:

$$Z_{0t} \text{ becomes } Z^{*}_{0t} = \Delta X_t$$

$$Z_{1t} \text{ becomes } Z^{*}_{1t} = (X^{'}_{t-1}, t)^{'}$$

$$Z_{2t} \text{ becomes } Z^{*}_{2t} = (\Delta X^{'}_{t-1}, \ldots, \Delta X^{'}_{t-k}, 1)^{'}$$

Statistical Analysis of I(1) models - (12)

Conclusion: **the matrix reduction problem has** $p_1 = p + 1$ **for dimension with** $\lambda_{p_1} = 0$

- **Limit laws**

➔ **Generally, the limit law of the JOHANSEN statistical test depends on the determinist terms, constrained or not.**

➔ **For big samples (about 400), the asymptotic distribution of the statistic is well known since the middle of the 90s and was tabulated by simulation. These standard critical values are available in statistical tables.**

➔ **For small samples, JOHANSEN proposed in 2002, a Bartlett correction which consists in estimating the VECM and in calculating a correction coefficient which is multiplied to the standard critical value.**

- Description of an arbitrage strategy
- First results

Description of an arbitrage strategy (1)

- **LAG choice**

The first problem to solve in order to work with the considered VECM is to determine the LAG of the model: k. In an article from 1999, Lütkepohl and Saïkkonen suggest to use the AIC (Akaïke Information Criteria). After having collected a few pieces of information, it appeared that Hurvich and Tsai had proposed in 1991 a corrected version of the AIC because it overestimates the real LAG.

With a constant, this AIC_c is an estimator of the expectancy of Kullback-Leibler, that is the distance between the sample and the considered VECM model.

The selection of k is to be done while minimizing the AIC_c for different values $k \in \{0;\ldots;p_{max}\}$.

Description of an arbitrage strategy (2)

For univariate time series, we have chosen p_{max} the same way as Fumio Hayashi does in his book « *Econometrics* » (2000), ie:

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

For multivariate time series, we have decided to take: $p_{max} = 6$

The implementation of the AIC_c with the sample $(X_0, \ldots, X_T)$ has to be done while doing the following regression:

Description of an arbitrage strategy (3)

$$Y = ZCoef + \varepsilon \text{ with } \mathrm{Y} = (\mathrm{X}_{\mathrm{p_{max}(T)+2}},\ldots,\mathrm{X}_{\mathrm{T}})' \ ,$$

$$\mathrm{Z} = \begin{pmatrix} X'_{\mathrm{p_{max}(T)+1}} & \Delta X'_{\mathrm{p_{max}(T)+1}} & \cdots \Delta X'_{\mathrm{p_{max}(T)+1-}k} & D'_1 \\ \vdots & \vdots & \vdots & \vdots \\ X'_{T-1} & \Delta X'_{T-1} & \cdots \quad X'_{T-1-k} & D'_T \end{pmatrix}$$

$$Coef = \begin{pmatrix} \Pi'_1 \\ \vdots \\ \Pi'_k \end{pmatrix}, \ \varepsilon \text{ with law } \mathrm{N_p}(0,\Sigma) \text{ and } \mathrm{X}'_0 = \ldots = X'_{-k+1} = 0$$

The OLS estimators give: $\hat{\mathrm{Coef}} = (Z'Z)^{-1} Z'Y$ and

$$\hat{\Sigma} = \frac{1}{T}(Y - Z\,\hat{Coef})'(Y - Z\,\hat{Coef})$$

Description of an arbitrage strategy (4)

In this framework:

$$AIC_c = T \ln\left|\hat{\Sigma}\right| + \frac{T(Tp+(kp+d)p)}{T-kp-d-p-1}$$

where d is the number of determinist terms.

- **<u>Estimation of the cointegration vectors:</u> it has to be implemented exactly the same way as it was described in Part I.**

- **<u>Sample size choice:</u> different backtestings showed that the concepts of stationarity and moreover of arbitrage are very furtive. So we decided to work on small samples, typically with a size of 50.**

Description of an arbitrage strategy (5)

❑ Stationarity test

In an article from March 2004 entitled *« Recent Advances in Cointegration Analysis »* , Helmut Lütkepohl advises to use the ADF (Augmented Dickey Fuller) test.

In case y_t is an AR(k), the ADF test uses the regression of Δy_t on $(y_{t-1}, \Delta y_{t-1}, \ldots, \Delta y_{t-k+1}, D_t^{'})$

and is based on the t-statistic $\hat{\tau}$ of the coefficient Π of the associated ECM model. Indeed:

$$\Delta y_t = \Pi y_{t-1} + \sum_{i=1}^{k-1} a_i \Delta y_{t-i} + \Phi D_t + \varepsilon_t \text{ ie } Y = Z\,Coef + \varepsilon \text{ with } Y = (\Delta y_{k+1}, \ldots, \Delta y_T)^{'} ,$$

$$Z = \begin{pmatrix} y_k & \Delta y_k & \cdots & \Delta y_2 & D_{k+1}^{'} \\ \vdots & \vdots & & \vdots & \vdots \\ y_{T-1} & \Delta y_{T-1} & \cdots & \Delta y_{T-k+1} & D_T^{'} \end{pmatrix}, \ \varepsilon_t \text{ with law } N(0,\sigma^2) , \ \hat{Coef} = (Z^{'}Z)^{-1} Z^{'} Y$$

$$\text{and } \hat{\sigma}^2 = \frac{\left\| Y - Z \hat{Coef} \right\|^2}{(T-k)-(k+d)} \text{ where } d = \dim(D_T)$$

Description of an arbitrage strategy (6)

$$\text{The test is: } \left| \begin{array}{ll} H_0: \quad \Pi = 0 & \text{ie } y_t \text{ non stationary} \\ H_1: \quad \Pi < 0 & \text{ie } y_t \text{ stationary} \end{array} \right.$$

$$\text{and } \hat{\tau} = \frac{\hat{\Pi}}{\sqrt{\hat{\sigma}^2 (Z'Z)^{-1}_{11}}} = \frac{\hat{Coef_1}}{\sqrt{\hat{\sigma}^2 (Z'Z)^{-1}_{11}}}$$

The limit law of this statistic is non standard and depends on the determinist terms of the model. The major drawback of this test is its lack of power for small samples (this is precisely the interesting case for us). So we decided to consider instead the ADF-GLS test, which is described in the book of Davidson&MacKinnon « Econometric Theory and Methods ».

Description of an arbitrage strategy (7)

This ADF-GLS was proposed by Elliott, Rothenberg and Stock in an article from 1996 entitled: « Efficient Tests for an Autoregressive Unit Root ».

While writing: $\Delta y_t = \gamma D_t + \beta y_{t-1} + \sum_{j=1}^{p} \delta_j \Delta y_{t-j} + \varepsilon_t$ **, the idea is to infer γ (ie the determinist coefficients) before infering β, because it appears on classical ADF tests, that the more determinist terms we have, the weaker the power of the test is.**

So, we consider the following regression:

$$y_t - \bar{\rho}\, y_{t-1} = (D_t - \bar{\rho}\, D_{t-1})\,\gamma^0 + v_t \text{ where } \bar{\rho} = 1 + \frac{\bar{c}}{T-1}$$

$$\text{with} \left| \begin{array}{l} \bar{c} = -7 \text{ when } D_t = (1) \\ \bar{c} = -13.5 \text{ when } D_t = (1, t) \end{array} \right.$$

Description of an arbitrage strategy (8)

Remark: $\overline{\rho} \underset{T \to +\infty}{\to} 1$

Let $\hat{\gamma}^0$ **be the estimator of** γ^0 **and let** $y_t^{'} = y_t - D_t \hat{\gamma}^0$

The regression: $\Delta y_t^{'} = \beta^{'} y_{t-1}^{'} + \sum_{j=1}^{p} \delta_j \Delta y_{t-j}^{'} + \varepsilon_t$ **helps us to calculate the t-statistic for** $\beta^{'} = 0$ **.**

If $D_t = (1)$ **then the asymptotic distribution of this statistic is** τ_{nc}

If $D_t = (1, t)$ **then the asymptotic distribution was tabulated by Elliott, Rothenberg and Stock, and is close to** τ_c

- **<u>JOHANSEN 's Rank Test</u>**

The implementation of this test comes from an article of JOHANSEN (2002) published in Econometrica and entitled: *« A small sample correction for the test of cointegration rank in the vector autoregressive model »*

Description of an arbitrage strategy (9)

Description of the chosen strategy

- The chosen strategy is simple.
- From a portfolio of p basic assets ($p \geq 10$), all the sub-portfolios of size 2,3 or 4 are extracted (meaning all pairs, triplets and quadruplets).
- For each sub-portfolio, we infer at instant t the vector $\hat{\beta}$ corresponding to the biggest eigenvalue of the associated VECM, in order to build a linear combination of the basic assets. This combination is called a synthetic asset.
- The stationarity at level 99% of the synthetic asset is checked in order to build an asset without trend, which is a modelling of the stochastic part of the synthetic asset.
- Therefore, this asset without trend is stationary around zero.
- Consequently, a good measure of the market risk of the synthetic asset is the standard deviation of the asset without trend.

Description of an arbitrage strategy (10)

- This measure will determine the quantity of synthetic asset to trade, as well as the conditions of opening and closing an arbitrage.
- Eventually, several backtestings have shown the need to use additional rules, called *consistency rules*, in order to get a good ratio of positive operations.
- The first rule is that for every proposed arbitrage, we decided to check the following condition: for every moving sub-sample of a certain size (typically 20) of the asset without trend, one should have 45% of the values above zero and 45% under zero. This condition is a translation of the fact that an asset that is stationary around zero is supposed to swing around zero.
- Last, when there are several arbitrage operations left, we decided to choose *the best one in a certain sense.* Our purpose is to calculate the mean for every moving subsample of a certain size (typically 20) and then to calculate the maximum of the absolute values of these means. *The best arbitrage operation is the one with the lowest maximum.*

Description of an arbitrage strategy (11)

Remarks:

➔ **This last condition was decided in order to avoid clusters of bad operations and to mutualize market risk over time. Moreover, this condition is synonymous with the fact that a process that is stationary around zero is supposed to have a mean close from zero.**

➔ **The spirit of the strategy is to be strict on the opening conditions of an operation because once this operation is released, whatever happens, the trader is charged for it.**

- **At instant t+1, a buying (resp. a selling) operation is released when the value of the asset without trend is in the bracket [-2.5σ;-σ] (resp. [σ;2.5σ]).**
- **The quantity of synthetic asset to be traded is defined as a percentage of the value of the portfolio normalized by σ.**

Description of an arbitrage strategy (12)

•Conversely, once an operation is launched, it is closed only when one of the three following conditions is satisfied:

- the target is completed, ie the asset without trend is positive (resp. negative) for a buying (resp. a selling) operation.
- the operation lasts more than 100 opened days (# 5 months).
- the operation generates losses bigger than 10σ per share of synthetic asset.

First results (1)

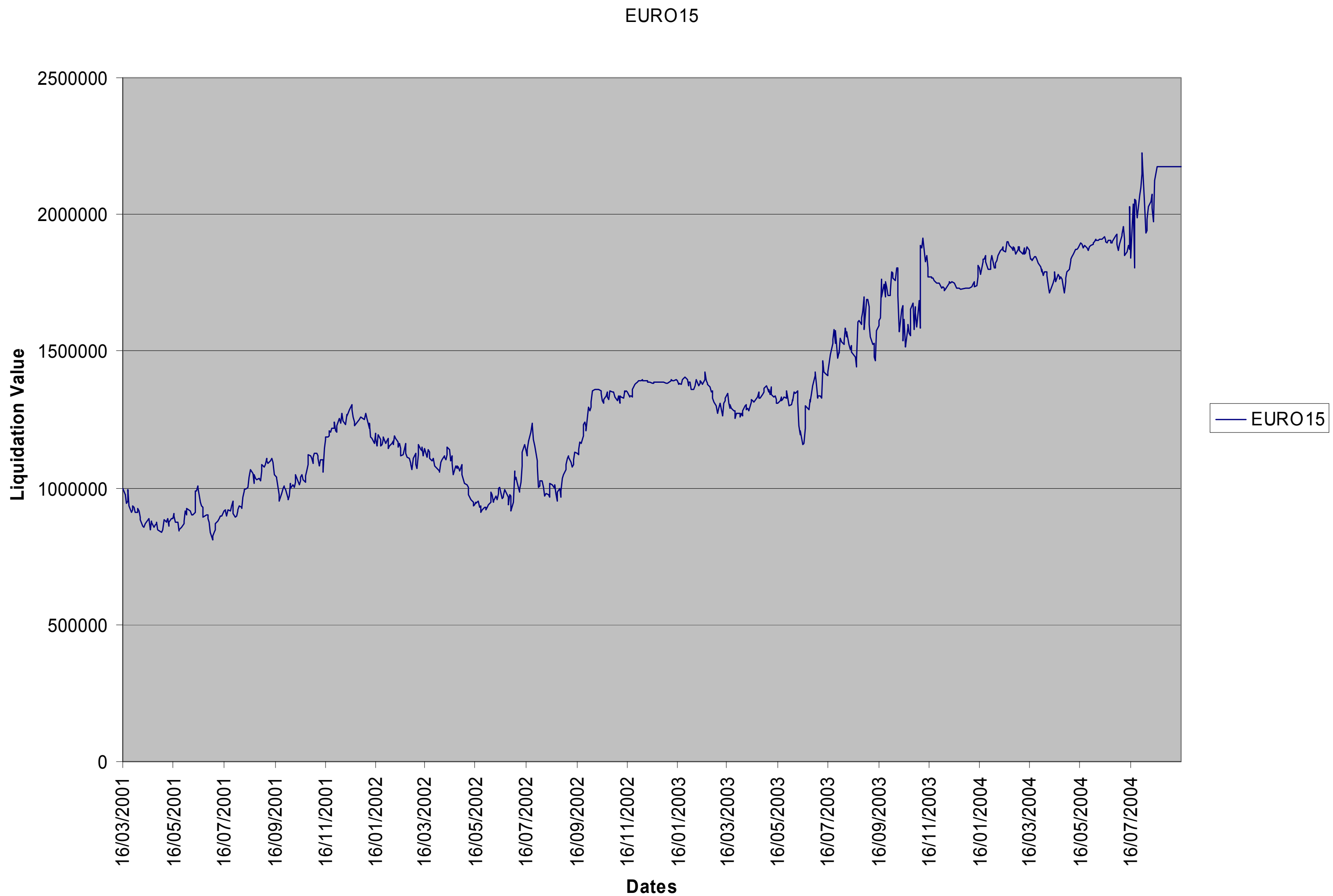
- The results we present are in no case definitive. They should be taken as an introduction to future developments.
- The following backtesting was made on the european market.
- <u>Choice of the data:</u>

  The single stocks used for this backtesting were these of 15 of the biggest capitalisations of the EuroStoxx50: ABN AMRO, Banco Bilbao Vizcaya Argentaria SA, Banco Santander Central Hispano SA, BNPParibas, Deutsche Bank AG, Deutsche Telekom AG, E.ON AG, ING Groep NV, Nokia OYJ, Royal Dutch Petroleum Co, Sanofi-Aventis, Siemens AG, Societe Generale, Telefonica SA et TOTAL SA.
- The study period begins in 03/16/2001 and ends on 09/14/2004. The prices used are the Last prices in Euros.
- Transaction cost are worth 10bp and the daily repo rate is taken at 4% (the biggest value from 2000 is about 3.5%)

First results (2)

- We obtained 99 operations opened and closed on the considered period. We backtested 105 pairs, 455 triplets et 1365 quadruplets.
- The Sharpe ratio is 3.67.
- The average P&L is 1.7σ, whereas the average length of an operation is 37 opened days.
- Last, the average annual growth rate is: 22%.
- The following graph describes the liquidation value of the portfolio.

First results (3)



Conclusion

- The strategy set up looks like a promising one.
- For equity single stocks, it would be interesting to work with portfolios greater or equal to 20 basic assets. This would generate IT problems, since the backtesting for 15 basic assets was nearly 3 days long. But we have begun the implementation of new libraries that should help us to divide by 2 or 3 the calculation time.
- It should be recalled that the only very important condition for such a strategy to work is the liquidity of the considered market.
- Therefore, our next step will be the application of this strategy to CMS rates. On one hand, we expect interesting results since CMS rates are very correlated, but on the other hand the way to valuate swaps is more complicated than for single stocks.