Econométrie des séries temporelles: Le modèle à correction d'erreurs (Japon)

Yipeng CHEN & Perrine DIAS (M1 IREF-FQA)

30/03/2021

Contents

Présentation du projet

Dans ce projet, nous cherchons la relation existante entre les flux nets d'investissements directs étrangers en % du PIB et le crédit intérieur fourni au secteur privé en % du PIB comme proxy du développement du secteur financier au Japon.

Commençons par estimer l'équation suivante soit par MCO soit par MCE :

$Y_t = \beta_0 + \beta_1 * X_t + \epsilon_t$

ou bien

$X_t = \beta_0 + \beta_1 * Y_t + \epsilon_t$

Avec Y_t: le crédit intérieur fourni au secteur privé en % du PIB comme proxy du développement du secteur financier

Avec X_t: les flux nets d'investissements directs étrangers en % du PIB

On fixe le risque de première espèce α à 5% ($\alpha = 0.05$)

Le modèle à correction d'erreurs (MCE)

«Rien n'est trop tôt, ni trop tard, tout est à sa place.»

Le modèle à correction d'erreurs (MCE) est un modèle utilisé dans l'étude de séries chronologiques multiples particulièrement utilisé pour les données dont les variables ont des tendances stochastiques communes sur le long terme. Cela s'appelle la cointégration. Il est coutumier de se servir du MCE lorsque nous souhaitons estimer les effets à court et à long terme d'une série chronologique sur une autre.

Le MCE regroupe un modèle statique et un modèle dynamique dont les équations sont les suivantes:

Modèle statique:

$\beta_1 \Delta x_t$

Modèle dynamique:

$\beta_2(y_{t-1} - x_{t-1})$

==> $\Delta y_t = \beta_1 \Delta x_t + \beta_2(y_{t-1} - x_{t-1})$

Démarche à suivre

Afin d'estimer l'équation $Y_t = \alpha_0 + \alpha_1 * X_t + \epsilon_t$:

Nous employons les MCO si Y_t et X_t sont stationnaires.

Si Y_t et X_t sont DS et cointégrées, alors nous employons le modèle à correction d'erreurs (MCE).

Si les deux séries sont DS mais pas cointégrées, alors on différencie les séries à l'ordre 1 pour les rendre stationnaires, donc nous estimons :

$\Delta Y_t = \alpha_0 + \alpha_1 \Delta X_t + \epsilon_t$

avec $\Delta Y_t = Y_t - Y_{t-1}$ et $\Delta X_t = X_t - X_{t-1}$

Si Y_t est stationnaire mais X_t est DS, alors on différencie X_t à l'ordre 1 et on estime:

$Y_t = \alpha_0 + \alpha_1 \Delta X_t + \epsilon_t$

avec $\Delta X_t = X_t - X_{t-1}$

Si Y_t est DS mais X_t est stationnaire, alors on différencie Y_t à l'ordre 1 et on estime:

$\Delta Y_t = \alpha_0 + \alpha_1 X_t + \epsilon_t$

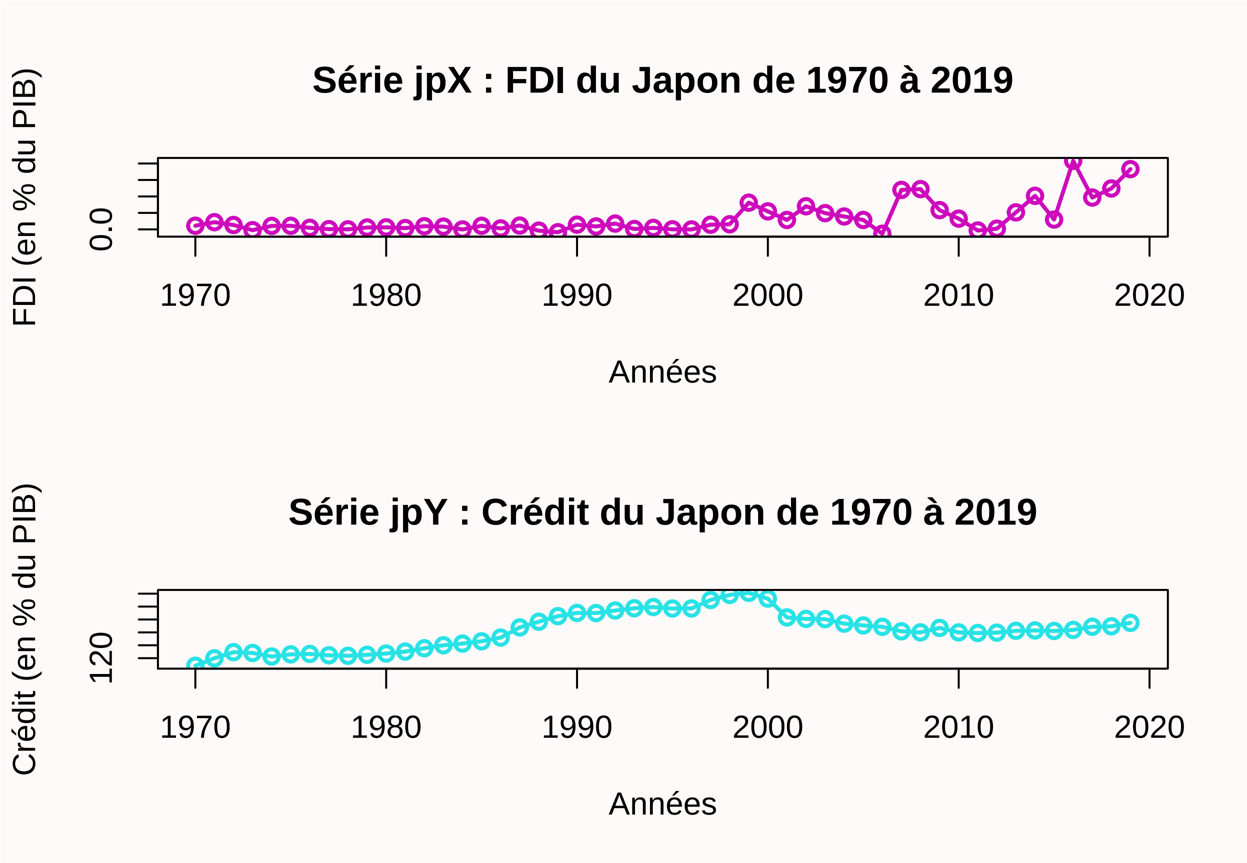
avec $\Delta Y_t = Y_t - Y_{t-1}$

Traitement des données

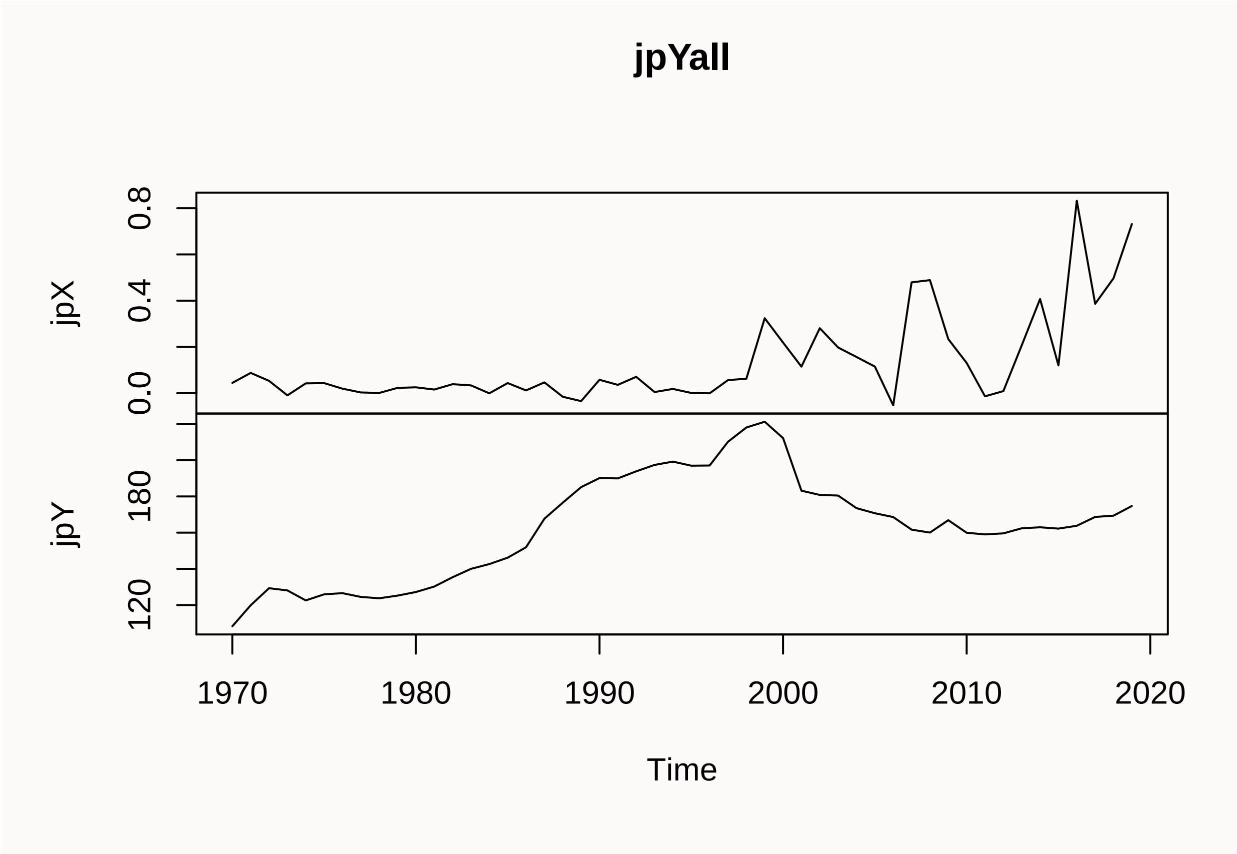
```r
library(readxl)

#==============================================================
#Credit
Credit<-readxl::read_excel("ADCtoPS.xls")

#Choisir les données relatives au Japon dans la table
jpY <- Credit[22,]
jpY=jpY[,2:ncol(jpY)]
jpY=as.numeric(jpY)
jpY=na.omit(jpY)
jpY=ts(jpY, freq=1, start=1970)
#=============================================================
#Flux investissements

fdi<-readxl::read_excel("NIofFDI.xls")

#Choisir les données relatives au Japon dans la table
jpX <- fdi[22,]
jpX=jpX[,2:ncol(jpX)]
jpX=as.numeric(jpX)
jpX=na.omit(jpX)
jpX=ts(jpX, freq=1, start=1970)

jpYall <- ts.union(jpX,jpY,dframe = FALSE)
```

Visualisation des deux séries

```r
op <- par(mfrow = c(2,1))
par(bg="snow1") #pour changer couleur du fond
plot.ts(jpX,xlab = "Années", ylab = "FDI (en % du PIB)", main = "Série jpX : FDI du Japon de 1970 à 2019
plot.ts(jpY,xlab = "Années", ylab = "Crédit (en % du PIB)", main = "Série jpY : Crédit du Japon de 1970
```




```
par(op)
plot(jpYall)
```



L'évolution est constante sur toute la période.

Tests de racine unitaire de LS sur les deux séries

```
###Test de LS pour la série temporelle "jpX"
source('C:/Users/perri/Bureau/M1/S8/Econométrie2/Séries_chrono/mce/LeeStrazicichUnitRootTestParalleliza
library(foreach)
library(doSNOW)
```

```
## Loading required package: iterators

## Loading required package: snow
```

```
library(parallel)
```

```
##
## Attaching package: 'parallel'

## The following objects are masked from 'package:snow':
##
##     clusterApply, clusterApplyLB, clusterCall, clusterEvalQ,
##     clusterExport, clusterMap, clusterSplit, makeCluster, parApply,
##     parCapply, parLapply, parRapply, parSapply, splitIndices,
##     stopCluster
```

```r
cl <- makeCluster(max(1, detectCores() - 1))
registerDoSNOW(cl)

#Avec une seule date de rupture || lag = 4 {lag = 5 ne marche pas}
myBreaks <- 1
myModel <- "break"
myLags <- 4
myParallel_LS <- ur.ls.bootstrap(y=jpX , model = myModel, breaks = myBreaks, lags = myLags, method = "F
```

```
## [[1]]
## [1] -3.661965
##
## [1] "First possible structural break at position: 37"
## [1] "The location of the first break - lambda_1: 0.7 , with the number of total observations: 50"
## Critical values - Break model:
##      lambda    1%    5%   10%
## [1,]    0.1 -5.11 -4.50 -4.21
## [2,]    0.2 -5.07 -4.47 -4.20
## [3,]    0.3 -5.15 -4.45 -4.18
## [4,]    0.4 -5.05 -4.50 -4.18
## [5,]    0.5 -5.11 -4.51 -4.17
## [1] "Number of lags used: 4"
## Runtime:
## Time difference of 0.003832984 mins
```

-3.661965 > -4.51 => On accepte H0 : $\phi = 0$

PGD DS avec 1 changement structurel en 2006

$X_t \sim I(1)$

Pour la série temporelle jpY:

```r
###Test de LS pour la série temporelle "jpY"
source('C:/Users/perri/Bureau/M1/S8/Econométrie2/Séries_chrono/mce/LeeStrazicichUnitRootTestParalleliza
library(foreach)
library(doSNOW)
library(parallel)
cl <- makeCluster(max(1, detectCores() - 1))
registerDoSNOW(cl)

#Avec une seule date de rupture || lag = 4 {lag = 5 ne marche pas}
myBreaks <- 1
myModel <- "break"
myLags <- 4
myParallel_LS <- ur.ls.bootstrap(y=jpY , model = myModel, breaks = myBreaks, lags = myLags, method = "F
```

```
## [[1]]
## [1] -2.900409
##
## [1] "First possible structural break at position: 31"
## [1] "The location of the first break - lambda_1: 0.6 , with the number of total observations: 50"
## Critical values - Break model:
```

```
##      lambda    1%    5%   10%
## [1,]    0.1 -5.11 -4.50 -4.21
## [2,]    0.2 -5.07 -4.47 -4.20
## [3,]    0.3 -5.15 -4.45 -4.18
## [4,]    0.4 -5.05 -4.50 -4.18
## [5,]    0.5 -5.11 -4.51 -4.17
## [1] "Number of lags used: 4"
## Runtime:
## Time difference of 0.003765265 mins
```

-2.900409 > -4.51 => On accepte H0 : $\phi = 0$

PGD DS avec 1 changement structurel en 2000

$Y_t \sim I(1)$

Les 2 séries sont I(1) = DS. On doit donc les stationnariser.

```
#Stationnariser les deux séries
djpX = diff (jpX)
djpY = diff (jpY)
```

A présent, nous rendons compte des étapes à suivre pour estimer la relation de long terme.

Les étapes à suivre afin de mettre en oeuvre le MCE dans le cadre bivarié:

1) Estimation par MCO d'une relation de long terme d'équation: $Y_t = \hat{\delta} + \hat{\gamma} X_t + \epsilon_t$
2) On vérifie que les résidus de cette régression sont I(0)
3) Estimation par les MCO de la relation de court terme (la relation du modèle dynamique):

$\Delta Y_t = \hat{\beta}_1 \Delta X_t + \hat{\beta}_2 \epsilon_{t-1} + u_t$

avec $\hat{\beta}_2 < 0$

et $\hat{\beta}_2$ qui mesure la force de rappel vers l'équilibre et doit être significativement négatif. Dans le cas contraire, on rejette une spécification MCE car le mécanisme de correction d'erreurs qui permet de tendre vers la relation de long terme divergerait.

Estimation de la relation de long terme

```
reg = lm(jpY~jpX)
summary(reg)
```

```
##
## Call:
## lm(formula = jpY ~ jpX)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -52.449 -23.996  -1.808  19.810  56.894
##
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  159.555      4.931  32.357   <2e-16 ***
## jpX           26.320     21.012   1.253    0.216
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 28.76 on 48 degrees of freedom
## Multiple R-squared:  0.03165,    Adjusted R-squared:  0.01148
## F-statistic: 1.569 on 1 and 48 DF,  p-value: 0.2164
```

γ est non significatif ici, car la p-value = 0.216 > 0.05 => On accepte H0 : $\gamma = 0$ => γ est statistiquement nul.

Or, on veut absolument que le γ soit significatif ! Par conséquent, on va enlever la constante δ dans la régression notée "reg".

On aura donc une nouvelle régression notée "reg1":

```r
reg1 = lm(jpY~jpX-1)
summary(reg1)
```

```
##
## Call:
## lm(formula = jpY ~ jpX - 1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -177.95   92.44  122.86  162.52  199.40
##
## Coefficients:
##     Estimate Std. Error t value Pr(>|t|)
## jpX   410.78      81.92   5.014 7.36e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 135.9 on 49 degrees of freedom
## Multiple R-squared:  0.3391, Adjusted R-squared:  0.3256
## F-statistic: 25.14 on 1 and 49 DF,  p-value: 7.364e-06
```

γ significatif ici car la p-value = 7.36e-06 < 0.05 => On rejette H0 : $\gamma = 0$ => γ est significativement différent de 0.

On vérifie que les résidus de cette régression sont I(0). Nous effectuons donc le test de LS pour les résidus de la régression notée "reg1".

On ne prend pas en compte l'autocorrélation dans les aléas de la régression notée "reg1" ici.

```r
#Avec une seule date de rupture
#lag = 0 {Sans modéliser l'autocorrélation dans les aléas de la régression notée <<reg1>>}
myBreaks <- 1
myModel <- "break"
```

```r
myLags <- 0
myParallel_LS <- ur.ls.bootstrap(y=reg1$res , model = myModel, breaks = myBreaks, lags = myLags, method
```

```
## [[1]]
## [1] -6.032144
##
## [1] "First possible structural break at position: 29"
## [1] "The location of the first break - lambda_1: 0.6 , with the number of total observations: 50"
## Critical values - Break model:
##      lambda    1%    5%   10%
## [1,]    0.1 -5.11 -4.50 -4.21
## [2,]    0.2 -5.07 -4.47 -4.20
## [3,]    0.3 -5.15 -4.45 -4.18
## [4,]    0.4 -5.05 -4.50 -4.18
## [5,]    0.5 -5.11 -4.51 -4.17
## [1] "Number of lags used: 0"
## Runtime:
## Time difference of 0.0009902318 mins
```

-6.032144 < -4.51 => On rejette H0 : $\phi = 0$ => Pas de Racine Unitaire. Les résidus de la régression notée "reg1" sont bien I(0).

Estimation de la relation de court terme

```r
n= length(jpX)
n
```

```
## [1] 50
```

```r
lres <- rep(0,n-1)# création du vecteur nul e(t-1) avant la boucle

for(i in 2:length(jpX)){lres[i]<-jpY[i-1]-410.78*jpX[i-1]}

lres<-lres[2:n] #e en (t-1) ; On enlève la première valeur qui doit être un NA
lres
```

```
##  [1]   90.108216   83.953532  107.489104  132.080732  105.257568  108.117752
##  [7]  118.647090  123.292199  123.414769  115.903115  116.877783  123.861661
## [13]  119.495012  126.256354  142.951433  128.300197  147.146810  148.617608
## [19]  183.007504  199.403850  166.475200  175.253242  164.916246  195.469283
## [25]  191.807365  196.675306  197.418311  187.078997  192.498847   88.201416
## [31]  122.438249  136.161760   65.478088   99.438423  109.316143  123.574296
## [37]  190.338571  -35.124043  -40.758453   70.888326  106.314182  164.698874
## [43]  155.988735   77.567177   -4.312585  113.080042 -177.947231    9.991430
## [49]  -34.612900
```

```r
newreg<-lm(djpY~djpX+lres-1)
summary(newreg)
```

```
##
## Call:
## lm(formula = djpY ~ djpX + lres - 1)
##
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -30.2447  -1.9257   0.6043   3.4921  14.4105
##
## Coefficients:
##       Estimate Std. Error t value Pr(>|t|)
## djpX -0.117004   5.821497  -0.020    0.984
## lres  0.009354   0.007588   1.233    0.224
##
## Residual standard error: 6.878 on 47 degrees of freedom
## Multiple R-squared:  0.03354,    Adjusted R-squared:  -0.007582
## F-statistic: 0.8156 on 2 and 47 DF,  p-value: 0.4485
```

β_1 est non significatif. Le coefficient associé à lres β_2 n'est pas significatif et est positif. Par conséquent le vecteur cointégrant noté $\beta = \binom{\beta_1}{\beta_2}$ est nul.

On est donc obligé de rejeter la spécification du MCE dans ce cas. Il semble que nos 2 séries Y_t et X_t ne sont pas cointégrées.

Conclusion intermédiaire:

$X_t \sim I(1) = DS$

$Y_t \sim I(1) = DS$

Les 2 séries sont I(1) = DS mais elles ne sont pas cointégrées. Afin de palier ce soucis, nous proposons une nouvelle régression avec les deux séries différenciées.

```r
regd = lm(djpY~djpX)
summary(regd)
```

```
##
## Call:
## lm(formula = djpY ~ djpX)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -30.2942  -2.4891   0.1037   3.5466  14.4004
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.3385     0.9823   1.363    0.179
## djpX          1.2622     5.5922   0.226    0.822
##
## Residual standard error: 6.854 on 47 degrees of freedom
## Multiple R-squared:  0.001083,   Adjusted R-squared:  -0.02017
## F-statistic: 0.05094 on 1 and 47 DF,  p-value: 0.8224
```

L'équation associée à cette régression est:

$\Delta Y_t = \alpha_0 + \alpha_1 \Delta X_t + \epsilon_t$

α_0 et α_1 ne sont pas significatifs car leurs p-values sont toutes les deux supérieures à 0.05. On n'arrive pas à trouver une régression qui nous permet de conclure sur la relation entre X_t et Y_t.

Estimation inversée

Il nous reste encore une possibilité à tester afin de trouver la relation existante entre X_t et Y_t; c'est-à-dire de tester la relation suivante:

$X_t = \beta_0 + \beta_1 Y_t + \epsilon_t$

Commençons par l'estimation de la relation de long terme:

```
###Estimation de la relation de Long-terme
reg_1 = lm(jpX~jpY)
summary(reg_1)
```

```
##
## Call:
## lm(formula = jpX ~ jpY)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.19398 -0.10861 -0.06450  0.02284  0.69835
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.0633852  0.1589408  -0.399    0.692
## jpY          0.0012027  0.0009601   1.253    0.216
##
## Residual standard error: 0.1944 on 48 degrees of freedom
## Multiple R-squared:  0.03165,    Adjusted R-squared:  0.01148
## F-statistic: 1.569 on 1 and 48 DF,  p-value: 0.2164
```

γ est non significatif ici car la p-value = 0.216 > 0.05 => On accepte H0 : $\gamma = 0$

=> γ est statistiquement nul

Or, on veut absolument que le γ soit significatif ! Par conséquent, on va enlever la constante δ dans la régression notée "reg_1". On va donc avoir une nouvelle régression notée "reg___1":

```
reg__1 = lm(jpX~jpY-1)
summary(reg__1)
```

```
##
## Call:
## lm(formula = jpX ~ jpY - 1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.19210 -0.11678 -0.07851  0.03585  0.69674
##
## Coefficients:
##      Estimate Std. Error t value Pr(>|t|)
## jpY 0.0008255  0.0001646   5.014 7.36e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.1927 on 49 degrees of freedom
## Multiple R-squared:  0.3391, Adjusted R-squared:  0.3256
## F-statistic: 25.14 on 1 and 49 DF,  p-value: 7.364e-06
```

γ est significatif ici car la p-value = 7.36e-06 < 0.05 => On rejette H0 : $\gamma = 0$

=> γ est donc significativement différent de 0.

On vérifie que les résidus de cette régression sont I(0). On effectue donc le test de LS pour les résidus de la régression notée "reg___1“. On prend en compte l'autocorrélation dans les aléas de la régression notée”reg___1".

```
#Avec une seule date de rupture
#lag = 1 {Pour modéliser l'autocorrélation dans les aléas de la régression notée <<reg__1>>}
myBreaks <- 1
myModel <- "break"
myLags <- 1
myParallel_LS <- ur.ls.bootstrap(y=reg__1$res , model = myModel, breaks = myBreaks, lags = myLags, meth
```

```
## [[1]]
## [1] -4.605744
##
## [1] "First possible structural break at position: 44"
## [1] "The location of the first break - lambda_1: 0.9 , with the number of total observations: 50"
## Critical values - Break model:
##      lambda    1%    5%   10%
## [1,]    0.1 -5.11 -4.50 -4.21
## [2,]    0.2 -5.07 -4.47 -4.20
## [3,]    0.3 -5.15 -4.45 -4.18
## [4,]    0.4 -5.05 -4.50 -4.18
## [5,]    0.5 -5.11 -4.51 -4.17
## [1] "Number of lags used: 1"
## Runtime:
## Time difference of 0.001107649 mins
```

-4.605744 < -4.51 => On rejette H0 : $\phi = 0$

=> Il n'y a pas de racine unitaire. Les résidus de la régression notée "reg___1" sont bien I(0). Nous pouvons donc estimer la relation de court terme via le code suivant:

```
n= length(jpY)
n
```

```
## [1] 50
```

```
lres <- rep(0,n-1)# création du vecteur nul e(t-1) avant la boucle

for(i in 2:length(jpY)){lres[i]<-jpX[i-1]-0.0008255*jpY[i-1]}

lres<-lres[2:n] #e en (t-1) ; On enlève la première valeur qui doit être un NA
lres
```

```
##  [1] -0.04516419 -0.01151135 -0.05361235 -0.11545684 -0.05905551 -0.06061235
##  [7] -0.08520235 -0.09985385 -0.10135727 -0.08070612 -0.07986122 -0.09200076
```

```
## [13] -0.07306967 -0.08211172 -0.11850771 -0.07715341 -0.11377306 -0.09196064
## [19] -0.16148211 -0.18753339 -0.09932497 -0.12099200 -0.08954664 -0.15823747
## [25] -0.14637184 -0.16185258 -0.16349116 -0.11720342 -0.11773746  0.14131263
## [31]  0.04345498 -0.03675177  0.13156304  0.04831312  0.01308404 -0.02613348
## [37] -0.19209112  0.34561172  0.35668865  0.09594350 -0.00148750 -0.14508996
## [43] -0.12294126  0.07246855  0.27269758 -0.01426758  0.69674723  0.24707928
## [49]  0.35678099
```

```
newreg<-lm(djpX~djpY+lres-1)
summary(newreg)
```

```
##
## Call:
## lm(formula = djpX ~ djpY + lres - 1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.17877 -0.06612 -0.03621 -0.00312  0.70657
##
## Coefficients:
##        Estimate Std. Error t value Pr(>|t|)
## djpY  3.465e-05  3.461e-03   0.010  0.99205
## lres -3.989e-01  1.367e-01  -2.918  0.00538 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1648 on 47 degrees of freedom
## Multiple R-squared:  0.1554, Adjusted R-squared:  0.1194
## F-statistic: 4.323 on 2 and 47 DF,  p-value: 0.01891
```

β_1 est non significatif donc β_1 est statistiquement nul.

Le coefficient associé à lres β_2 qui est significatif et négatif valide le MCE.

Le vecteur cointégrant noté $\beta = \binom{\beta_1}{\beta_2} = \binom{0}{-0.3989}$ est non nul.

Par conséquent, nos 2 séries X_t et Y_t sont en effet cointégrées et on a un comportement à court terme et à long terme.

Conclusion finale

Relation de long terme estimée :

$X_t = 0.0008255 Y_t$

$FDI_t = 0.0008255 * Crédit_t$

Relation de court terme estimée :

$\Delta X_t = -0.3989 e_{t-1}$

$\Delta FDI_t = -0.3989 e_{t-1}$

-0.3989 < 0 et est statistiquement significatif

Par conséquent, notre modélisation du MCE est validée.

Vérification de notre modèle à correction d'erreurs

En prenant l'année de 1971, on a :

Relation de long terme estimée :

$FDI_{1971} = 0.0008255 * Crédit_{1971}$

La valeur estimée de FDI en 1971 :

```
FDI_1971 = 0.0008255 * jpY[2]
FDI_1971
```

```
## [1] 0.09895604
```

La vraie valeur de FDI en 1971 :

```
FDI1971 = jpX[2]
FDI1971
```

```
## [1] 0.08744469
```

L'écart entre la valeur estimée et la vraie valeur :

```
ecartFDI = abs(FDI_1971 - FDI1971)
ecartFDI
```

```
## [1] 0.01151135
```

ecartFDI $= 0.01151135 < 0.02$

Relation de court terme estimée :

$\Delta FDI_{1971} = -0.3989 e_{1970}$

La valeur estimée de dFDI en 1971 :

```
dFDI_1971 = - 0.3989 * lres[1]
dFDI_1971
```

```
## [1] 0.018016
```

La vraie valeur de dFDI en 1971 :

```
dFDI1971 = djpX[1]
dFDI1971
```

```
## [1] 0.04323211
```

L'écart entre la valeur estimée et la vraie valeur :

```
ecartdFDI = abs(dFDI_1971 - dFDI1971)
ecartdFDI
```

```
## [1] 0.02521612
```

ecartdFDI $= 0.02521612 < 0.03$

L'erreur de l'estimation par MCE est assez faible !

On peut bien avoir confiance en notre Modèle à Correction d'Erreurs.

Annexes

Régression fallacieuse

Soit $y_t = y_{t-1} + u_t$ et $x_t = x_{t-1} + v_t$ avec u_t et v_t des bruits blancs. Si l'on effectue la régression de l'une sur l'autre de la manière suivante:

$y_t = \beta_0 + \beta_1 x_t + w_t$

avec w_t un bruit blanc, alors on obtient une régression fallacieuse: β_1 sera significatif même si y_t ne dépend pas de x_t. Cette régression n'a pas lieu d'être étant donné que y et x sont indépendants. En effet, y ne dépend que de sa valeur passée d'une période. La statistique t rejettera H0: $\beta_1 = 0$ trop souvent alors qu'elle est vraie selon *Granger et Newbold (1974)*. Asymptotiquement (quand T tend vers + ∞), t rejettera toujours H0 (*Phillips (1986)*).

Expérience de Monte Carlo (Davidson et MacKinnon (1993))

Nous avons toujours $y_t = y_{t-1} + u_t$ et $x_t = x_{t-1} + v_t$

Nous considérons les deux régressions suivantes sachant qu'elles ont été réalisées sur 10000 échantillons et on comptabilise la proportion de rejets de H0: $\alpha_1 = 0$ avec un seuil de 5% sachant que si la théorie asymptotique classique s'appliquait, ce taux de rejet devrait être de 5%.

	A	B
T	rejet de $H_0 : \alpha_1 = 0$	rejet $H_0 : \alpha_1 = 0$
25	0.53	0.15
50	0.66	0.15
100	0.76	0.16
500	0.89	0.17
1000	0.93	0.17

Figure 1: .

Régression A:

$y_t = \alpha_0 + \alpha_1 x_t + \epsilon_t$

Régression B:

$y_t = \alpha_0 + \alpha_1 x_t + \alpha_2 y_{t-1} + \epsilon_t$

Sur le tableau précédent (*Figure 1*), la colonne A représente la proportion de rejets de H0 alors qu'elle est vraie dans la régression A. La colonne B rend compte de la proportion de rejets de H0 alors qu'elle est vraie dans la régression B.

La cointégration dans le cadre bivarié

Deux séries x_t et y_t sont dites cointégrées si:

1) Elles sont intégrées du même ordre
2) Il existe un vecteur $\beta = \binom{\beta_1}{\beta_2}$ non nul tel que:

$\beta_1 x_t + \beta_2 y_t$ est une série d'ordre d'intégration inférieur

On suppose que $y_t \sim I(1)$, $x_t \sim I(1)$ et qu'il existe un vecteur cointégrant noté β entre les deux séries.

Une régression directe de y_t sur x_t mettrait simplement en évidence une relation entre les deux tendances des séries et non pas entre les séries.

Le problème est donc de retirer la relation commune de cointégration (la tendance commune) et aussi de rechercher la relation réelle entre les variables.

Solution: MCE