Econométrie des séries temporelles Projet 2022 - Taux d'intérêt de court terme de l'Allemagne

Dulymamode M1 IREF

24 Mars 2022

Sommaire

1 Description des données

Nous disposons des taux d'intérêts annuels de court terme allemand entre 1960 et 2021. Notre échantillon de données est donc constitué de 62 observations. Le but est de construire un modèle afin d'obtenir de l'information sur le processus qui a engendré les données appelé Processus Générateur de Données (PGD). Ceci sera fait sur le logiciel RStudio.

Nous allons construire la série temporelle ict (taux d'intérêt à court terme):

Code:

```
ict<-ts(data, start=c(1960,1), freq=1)
```

Nous commençons par une visualisation des données en utilisant un chronogramme. On observe une tendance décroissante sur l'ensemble des données. Il n'y a pas de saisonnalité car nous sommes sur des données annuelles.

Code:

```
plot.ts(ict,xlab='année', ylab='taux d'intérêt de court terme', main="Taux d'intérêt de court terme de l'Allemagne",col=1,type="o")
m=mean(ict)
abline(h=m,col=2)
```

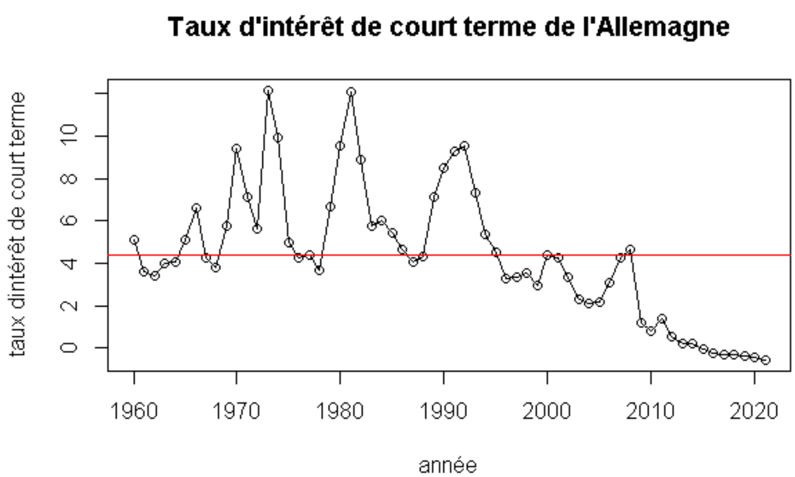

Figure 1: Evolution du taux d'intérêt de court terme

2 Tests de Racine Unitaire

Pour modéliser le PGD, il faut d'abord déterminer si la série temporelle est stationnaire ou pas. Les valeurs d'une série stationnaire varient dans le temps mais la manière dont elles varient ne change pas. Par contre pour une série non stationnaire, la manière dont les valeurs varient change dans le temps. Pour cela nous allons utiliser les tests de Racine unitaire. Les tests de racine unitaire permettent de déterminer si le PDG est stationnaire, TS (Trend-Stationary) ou DS (Difference-Stationary). Les quatres tests de racine unitaire utilisés dans le cadre de ce projet sont:

- Test de Dickey Fuller (DS)
- Test de Dickey Fuller Augmenté (ADF)
- Test de Zivot et Andrews (ZA)
- Test de Lee et Strazicich (LS)

2.1 Test de Dickey Fuller (DF)

Le test de Dickey-Fuller est un test à racine unique qui détecte statistiquement la présence d'un comportement de tendance stochastique dans la série chronologique des variables au moyen d'un test d'hypothèse. Pour DF, il existe 3 spécifications possibles: "trend", "drift" ou "none". Si la spécification est "trend" nous avons l'équation suivante:

$$\triangle X_t = (\rho - 1)X_{t-1} + \beta_0 + \beta_1 tendance_t + \epsilon_t \tag{1}$$

on teste les hypothèses suivantes:

$$H_0 : \rho - 1 = 0, \beta_1 = 0 \tag{2}$$

$$H_a : |\rho| < 1, \beta_1 \neq 0 \tag{3}$$

```
> summary(ur.df(ict,type="trend",lag=0))

###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression trend


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.9225 -0.8136 -0.3395  0.4092  6.2819 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)   
(Intercept)  2.29198    0.82452   2.780  0.00732 **
z.lag.1     -0.27896    0.08975  -3.108  0.00291 **
tt          -0.03669    0.01559  -2.353  0.02205 * 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.693 on 58 degrees of freedom
Multiple R-squared:  0.1468,	Adjusted R-squared:  0.1174 
F-statistic: 4.989 on 2 and 58 DF,  p-value: 0.01001


Value of test-statistic is: -3.1083 3.3871 4.9894 

Critical values for test statistics: 
      1pct  5pct 10pct
tau3 -4.04 -3.45 -3.15
phi2  6.50  4.88  4.16
phi3  8.73  6.49  5.47
```

Figure 2: Résultat de DF

Pour tester si $\beta_1 = 0$, il faut regarder la p-value de β_1. Ici la p-value est de $0.02205 < 0.05$. Donc avec un risque de première espèce de 0.05, on trouve que le coefficient β_1 est significatif. On regarde la statistique t associée à $\rho - 1$. On a -3.108>-3.45. Comme elle est supérieure à la valeur qui se trouve à l'intersection de la colonne 5% et ligne tau3, on en déduit que la spécification "Trend" est la bonne. Nous avons $\beta_1 \neq 0$ et que $\rho - 1 = 0$ alors notre série est DS.

On peut se référer au graphique suivant pour savoir si notre série est DS, TS ou stationnaire.

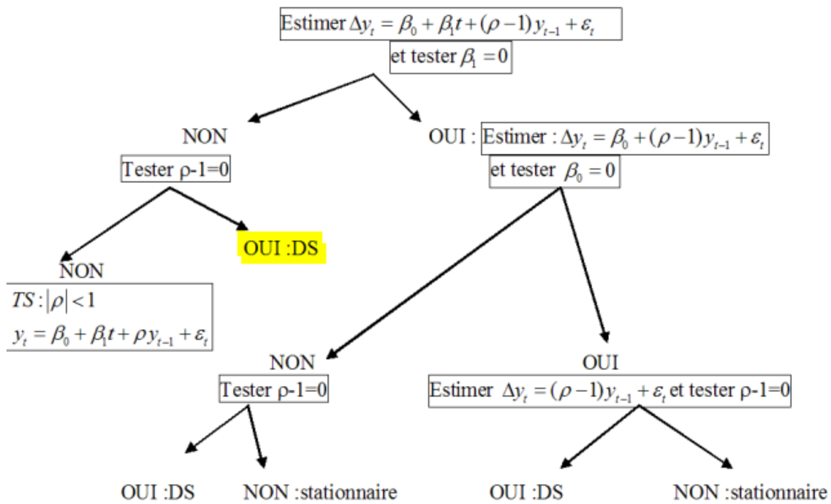

Figure 3: Schéma série DS, TS, stationnaire

Cependant les conclusions du test de Dickey Fuller ne sont valables que si les ϵ_t ne sont pas autocorrélés. On teste cela avec l'ACF et le PACF.

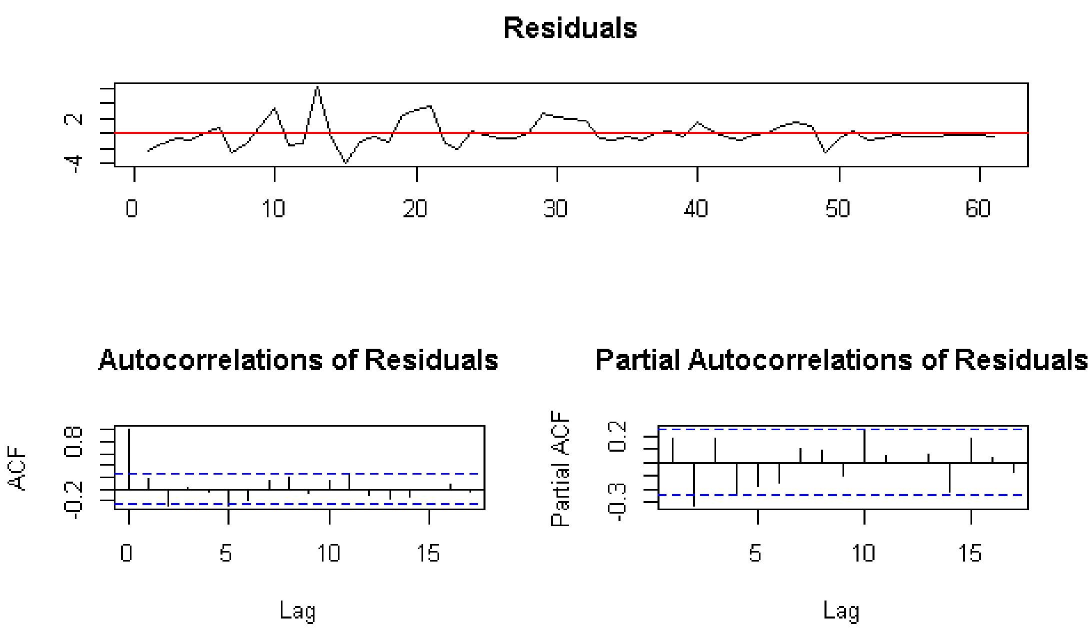

Figure 4: ACF, PACF

On remarque qu'il y a de l'autocorrélation, donc il faut passer au test de Dickey Fuller Augmenté (ADF).

2.2 Test de Dickey Fuller Augmenté (ADF)

L'ADF ajoute des variables explicatives en plus jusqu'à l'ordre p. Comme on garde généralement la même spécification que pour DF. On peut s'attendre à une équation de la forme suivante:

$$\triangle X_t = (\rho - 1)X_{t-1} + \beta_0 + \beta_1 t + \epsilon_t + \sum_{p=1}^{P} \gamma_p \triangle X_{t-p} \tag{4}$$

2.2.1 MAIC

Afin de déterminer le nombre de p retards qu'il faut ajouter, on utilise le MAIC. Voici la formule du MAIC avec p retards:

$$MAIC(P) = ln(\widehat{\sigma_p^2}) + 2\frac{\tau_T(p) + p}{T - pmax} \tag{5}$$

$$\tau_T(p) = \frac{\widehat{\sigma_0^2} \sum_{t=pmax+1}^{T} y_{t-1}}{\widehat{\sigma_p^2}} \tag{6}$$

$$\widehat{\sigma_p^2} = \frac{1}{(T - pmax)} * \sum_{t=pmax+1}^{T} e_t^2 \tag{7}$$

avec σ_0^2 qui provient de l'estimation par les MCO de la régression suivante:

$$\triangle \tilde{y}_t = \alpha_0 \tilde{y}_{t-1} + \sum_{p=1}^{P} \alpha_p \triangle \tilde{y}_{t-p} + \epsilon_t \tag{8}$$

ϵ_t sont les résidus et

$$\tilde{y}_t = y_t - \widehat{\beta\,'} z_t \tag{9}$$

```
> #MAIC
> summary(CADFtest(ict,criterion="MAIC",type="trend",max.lag.y=pmax))
Augmented DF test
                                              ADF test
t-test statistic:                          -3.51748935
p-value:                                    0.04809759
Max lag of the diff. dependent variable:    0.00000000

Call:
dynlm(formula = formula(model), start = obs.1, end = obs.T)

Residuals:
    Min      1Q  Median      3Q     Max
-4.0592 -0.7276 -0.2655  0.3988  5.5109

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.34863    1.43867   3.023  0.00401 **
trnd        -0.07356    0.02624  -2.804  0.00727 **
L(y, 1)     -0.40958    0.11644  -3.517  0.04810 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.651 on 48 degrees of freedom
Multiple R-squared:  0.2049,    Adjusted R-squared:  0.1718
F-statistic:    NA on NA and NA DF,  p-value: NA
```

Figure 5: MAIC

Comme le MAIC donne 0, on utilise le BIC

2.2.2 BIC

$$BIC = -2log(L) + k * log(T) \tag{10}$$

avec

- L: La vraisemblance du modèle estimé
- T: Le nombre d'observations dans l'échantillon
- k: Le nombre de paramètres libres dans le modèle

```
> summary(ur.df(ict,type="trend",lag=pmax,selectlag="BIC"))

###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression trend


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
    Min      1Q  Median      3Q     Max
-3.1688 -0.6095 -0.1523  0.7217  4.1448

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.47157    1.62084   3.993 0.000239 ***
z.lag.1     -0.62676    0.14491  -4.325 8.36e-05 ***
tt          -0.10531    0.02778  -3.791 0.000444 ***
z.diff.lag1  0.52597    0.14379   3.658 0.000664 ***
z.diff.lag2 -0.12260    0.12743  -0.962 0.341157
z.diff.lag3  0.36809    0.12786   2.879 0.006087 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.451 on 45 degrees of freedom
Multiple R-squared:  0.4244,    Adjusted R-squared:  0.3605
F-statistic: 6.637 on 5 and 45 DF,  p-value: 0.0001049


Value of test-statistic is: -4.3252 6.5336 9.3596

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -4.04 -3.45 -3.15
phi2  6.50  4.88  4.16
phi3  8.73  6.49  5.47
```

Figure 6: BIC

On trouve que:

- γ_3: $|2.876| > 1.64$
- γ_2: $|$-0.962$| < 1.64$
- γ_1: $|3.658| > 1.64$

Le BIC ajoute γ_1, γ_2 et γ_3 pour prendre en compte l'autocorrélation.

On peut aussi utiliser une autre méthode pour définir le nombre de γ_p. On ajoute un lag de 7 au modèle puis on enlève les γ non significatifs. Pour déterminer cela on regarde si leur statistique t en valeur aboslue est supérieure à 1,64.

```
> summary(ur.df(ict, type ="trend",lag=7))

###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression trend


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
    Min      1Q  Median      3Q     Max
-2.3451 -0.7439 -0.0425  0.6651  4.2775

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.33891    1.39725   3.105  0.00332 **
z.lag.1     -0.39144    0.14552  -2.690  0.01006 *
tt          -0.07871    0.02247  -3.504  0.00107 **
z.diff.lag1  0.29447    0.17132   1.719  0.09268 .
z.diff.lag2 -0.32529    0.16006  -2.032  0.04819 *
z.diff.lag3  0.16853    0.16046   1.050  0.29932
z.diff.lag4 -0.20965    0.14861  -1.411  0.16535
z.diff.lag5 -0.16265    0.15239  -1.067  0.29165
z.diff.lag6 -0.28268    0.12874  -2.196  0.03343 *
z.diff.lag7 -0.05649    0.13503  -0.418  0.67773
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.436 on 44 degrees of freedom
Multiple R-squared:  0.5053,    Adjusted R-squared:  0.4042
F-statistic: 4.994 on 9 and 44 DF,  p-value: 0.0001211


Value of test-statistic is: -2.6899 4.4169 6.1492

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -4.04 -3.45 -3.15
phi2  6.50  4.88  4.16
phi3  8.73  6.49  5.47
```

Figure 7: DF avec un lag de 7

On trouve que la statistique t de γ_7 en valeur absolue (-0.418) est inférieure à 1.64. On enlève γ_7.

```
> summary(ur.df(ict, type="trend",lag=6))

###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression trend


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
    Min      1Q  Median      3Q     Max
-3.3692 -0.7987 -0.0361  0.5381  4.5426

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.09830    1.36593   2.268   0.0281 *
z.lag.1     -0.29331    0.14462  -2.028   0.0484 *
tt          -0.05765    0.02202  -2.618   0.0119 *
z.diff.lag1  0.23063    0.16324   1.413   0.1644
z.diff.lag2 -0.38887    0.15727  -2.473   0.0172 *
z.diff.lag3  0.14803    0.15378   0.963   0.3408
z.diff.lag4 -0.25496    0.15475  -1.648   0.1063
z.diff.lag5 -0.14971    0.13314  -1.124   0.2666
z.diff.lag6 -0.28048    0.13529  -2.073   0.0438 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.509 on 46 degrees of freedom
Multiple R-squared:  0.4442,    Adjusted R-squared:  0.3476
F-statistic: 4.596 on 8 and 46 DF,  p-value: 0.0003625


Value of test-statistic is: -2.0281 2.6967 3.4286

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -4.04 -3.45 -3.15
phi2  6.50  4.88  4.16
phi3  8.73  6.49  5.47
```

Figure 8: DF avec un lag de 6

On retient le modèle avec un lag de 6 car (γ_6) est significatif (|-2.073| > 1.64). La spécification trend est la bonne comme la p-value de β_1 (0.0119) est inférieure à 0.05. On accepte $H_0 : \rho - 1 = 0$ car la statistique t de $\rho - 1$ (-2.028) est plus grande que -3.45. Notre processus est alors DS (Difference stationary) avec présence de racine unitaire.

2.3 Test de Zivot et Andrews

Les changements structurels ne sont pas détectés par les tests de DF et d'ADF. Ces changements peuvent toucher le niveau de la série, ses autocorrélations et sa variance. Pour détecter ces changements nous allons utiliser le test de Zivot et Andrews (1992).

Le modèle avec la spécification "intercept":

$$y_t = \beta_0 + \beta_1 t + \rho y_{t-1} + \delta DU_t(T_B) + \sum_{j=1}^{P} \gamma_j \triangle y_{t-j} + \epsilon_t \tag{11}$$

Le modèle avec la spécification "both":

$$y_t = \beta_0 + \beta_1 t + \rho y_{t-1} + \delta_1 DU_t(T_B) + \delta_2 DT_t(T_B) + \sum_{j=1}^{P} \gamma_j \triangle y_{t-j} + \epsilon_t \tag{12}$$

on teste les hypothèses suivantes:

$$H_0 : \text{DS sans changement structurel, } \rho = 1 \tag{13}$$

$$H_a : \text{TS avec unique changement structurel, } |\rho| < 1 \tag{14}$$

Il existe 3 spécifications pour ce test notamment, "both", "trend" et intercept. Sachant que le modèle "changing trend" n'est pas adapté pour les séries économiques, on aura pas la spécification trend. Pour débuter on commence par la spécification "both" pour voir si les δ sont significatifs.

La valeur du lag est obtenue par le BIC.

```
> summary(ur.za(ict, model="both",lag=3))

################################
# Zivot-Andrews Unit Root Test #
################################


Call:
lm(formula = testmat)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.5503 -0.6078 -0.0888  0.5375  5.2911 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  3.53231    1.13032   3.125  0.00296 ** 
y.l1         0.16807    0.14823   1.134  0.26226    
trend        0.12784    0.08253   1.549  0.12769    
y.dl1        0.65814    0.14073   4.677 2.25e-05 ***
y.dl2       -0.02658    0.12446  -0.214  0.83178    
y.dl3        0.42482    0.12284   3.458  0.00112 ** 
du           0.95649    0.81914   1.168  0.24848    
dt          -0.30613    0.09920  -3.086  0.00330 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.374 on 50 degrees of freedom
  (4 observations deleted due to missingness)
Multiple R-squared:  0.8409,	Adjusted R-squared:  0.8187 
F-statistic: 37.76 on 7 and 50 DF,  p-value: < 2.2e-16


Teststatistic: -5.6123 
Critical values: 0.01= -5.57 0.05= -5.08 0.1= -4.82 

Potential break point at position: 20 
```

Figure 9: ZA avec la spécification 'both'

On trouve que δ_1 n'est pas significatif mais que δ_2 l'est. Comme on ne peut pas se retrouver avec un modèle ayant seulement DT (modèle changing growth), on essaie la spécification "intercept".

```
> summary(ur.za(ict, model="intercept",lag=3))

################################
# Zivot-Andrews Unit Root Test #
################################


Call:
lm(formula = testmat)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.3263 -0.5926 -0.0430  0.8016  3.8845 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  4.22984    0.96079   4.402 5.49e-05 ***
y.l1         0.31904    0.13427   2.376 0.021287 *  
trend       -0.11752    0.02502  -4.698 2.03e-05 ***
y.dl1        0.55169    0.13452   4.101 0.000148 ***
y.dl2       -0.12032    0.12070  -0.997 0.323531    
y.dl3        0.38371    0.12368   3.102 0.003127 ** 
du           3.07051    0.95995   3.199 0.002375 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.418 on 51 degrees of freedom
  (4 observations deleted due to missingness)
Multiple R-squared:  0.8271,	Adjusted R-squared:  0.8068 
F-statistic: 40.67 on 6 and 51 DF,  p-value: < 2.2e-16


Teststatistic: -5.0717 
Critical values: 0.01= -5.34 0.05= -4.8 0.1= -4.58 

Potential break point at position: 10 
```

Figure 10: ZA avec la spécification 'intercept'

Résultat:

- δ_1 a une p-value de 0.00237 $<$ 0.05 donc la spécification "intercept" est correcte
- γ_3 est significatif car 3.102 (statistique t) >1.64
- ρ: $|0.319| < 1$, donc on rejette H_0
- La statistique t calculée, -5.0717 $<$ -4.8 donc on rejette H_0

On a donc un TS avec un unique changement structurel. La date de rupture est à la 10 ème observation. Donc $T_B = 1969$.

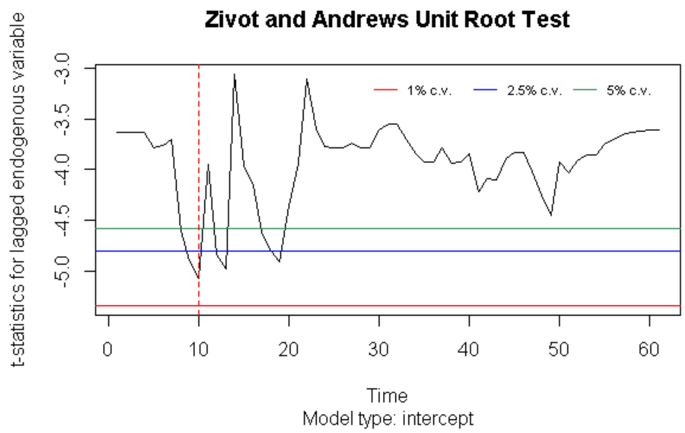

Figure 11: Zivot et Andrews test

2.4 Lee et Strazicich

Le problème avec ZA est qu'il ne peut y avoir qu'une seule date de rupture. Avec le test de Lee et Strazicich on peut avoir deux dates de rupture.

Pour la spécification "crash" nous avons les hypothèses suivantes:

$$H_0 : y_t = \mu_0 + d_1 B_1 t + d_2 B_2 t + y_{t-1} + v_1 t \tag{15}$$

$$H_1 : y_t = \mu_1 + \gamma * trend_t + d_1 D_1 t + d_2 D_2 t + v_2 t \tag{16}$$

Pour la spécification "break" nous avons les hypothèses suivantes:

$$H_0 : y_t = \mu_0 + d_1 B_1 t + d_2 B_2 t + d_3 D_1 t + d_4 D_2 t + y_{t-1} + v_1 t \tag{17}$$

$$H_1 : y_t = \mu_1 + \gamma * trend_t + d_1 D_1 t + d_2 D_2 t + d_3 D_1 t + d_4 D_2 t + v_2 t \tag{18}$$

Comme on avait seulement δ_1 significatif dans ZA, alors on choisit la spécification "crash". On définit une seule date de rupture pour le moment.

```
> myBreaks <- 1
> myModel <- "crash"
> myLags <- 3
> myLS_test <- ur.ls(y=ict , model = myModel, breaks = myBreaks, lags = myLags, method
 = "GTOS",pn = 0.1, print.results = "print" )
[1] -4.864436
[1] "First possible structural break at position: 13"
[1] "The location of the first break - lambda_1: 0.2 , with the number of total observa
tions: 62"
Critical values - Crash model:
        1%     5%    10%
[1,] -4.239 -3.566 -3.211
[1] "Number of lags determined by general-to-specific lag selection: 3"
Runtime:
Time difference of 0.01008493 mins
```

Figure 12: Lee et Strazicich test avec 1 date de rupture

Résultat:

- La valeur de la t statistique calculée est -4.8644
- λ est estimé à 0.2 et $T_b = 13$ donc 1972 est la date de rupture
- La valeur critique à 5% est -3.566
- Comme -4.8644 $<$ -3.566 donc on rejette H_0 donc le PGD est TS avec un changement structurel
- Le nombre de lag introduit dans le modèle pour tenir compte de l'autocorrélation est égal à 3

On teste de nouveau avec deux dates de rupture.

```
> myBreaks <- 2
> myLS_test <- ur.ls(y=ict , model = myModel, breaks = myBreaks, lags = myLags, method
 = "GTOS",pn = 0.1, print.results = "print" )
[1] -5.13523
[1] "First possible structural break at position: 13"
[1] "The location of the first break - lambda_1: 0.2 , with the number of total observa
tions: 62"
[1] "Second possible structural break at position: 18"
[1] "The location of the second break - lambda_2: 0.3 , with the number of total observ
ations: 62"
Critical values:
              Break 2 - 0.4 - 1% Break 2 - 0.4 - 5% Break 2 - 0.4 - 10%
Break 1 - 0.2              -6.16              -5.59               -5.27
Break 1 - 0.4                 NA                 NA                  NA
Break 1 - 0.6                 NA                 NA                  NA
              Break 2 - 0.6 - 1% Break 2 - 0.6 - 5% Break 2 - 0.6 - 10%
Break 1 - 0.2              -6.41              -5.74               -5.32
Break 1 - 0.4              -6.45              -5.67               -5.31
Break 1 - 0.6                 NA                 NA                  NA
              Break 2 - 0.8 - 1% Break 2 - 0.8 - 5% Break 2 - 0.8 - 10%
Break 1 - 0.2              -6.33              -5.71               -5.33
Break 1 - 0.4              -6.42              -5.65               -5.32
Break 1 - 0.6              -6.32              -5.73               -5.32
[1] "Number of lags determined by general-to-specific lag selection: 3"
Runtime:
Time difference of 0.07330972 mins
```

Figure 13: Lee et Strazicich test avec 2 dates de rupture

Résultat:

- La valeur de la t statistique calculée est -5.13523
- λ_1 est estimé à 0.2 et $T_b = 13$ donc 1972 est la date de rupture
- λ_2 est estimé à 0.3 et $T_b = 18$ donc 1977 est la date de rupture
- La valeur critique à l'intersection de $\lambda_1 = 0.2$ et $\lambda_2 = 0.3$ est -5.59
- Comme -5.13523 > -5.59, on ne rejette pas H_0

2.5 Bootstrap Lee et Strazicich

Comme nous avons un petit échantillon (62 observations), on utilise ls.bootstrap.

```
> cl <- makeCluster(max(1, detectCores() - 1))
> registerDoSNOW(cl)
> myBreaks <- 1
> myModel <- "crash"
> myModel <- "crash"
> myParallel_LS <- ur.ls.bootstrap(y=ict , model = myModel, breaks = myBreaks, lags = m
yLags, method = "Fixed",pn = 0.1, critval = "bootstrap", print.results = "print")
[[1]]
[1] -4.864436

[1] "First possible structural break at position: 13"
[1] "The location of the first break - lambda_1: 0.2 , with the number of total observa
tions: 62"
Critical values - Crash model:
        1%     5%    10%
[1,] -4.239 -3.566 -3.211
[1] "Number of lags used: 3"
Runtime:
Time difference of 0.03177096 mins
```

Figure 14: Lee et Strazicich test avec bootstrap

On compare la statistique t calculée à la valeur critique à 5%. Comme -4.864436 < -3.566, on rejette H_0. Le PGD qui a généré notre série est TS avec une date de rupture dans la constante et dans la tendance en 1969.

3 Série TS (Trend-Stationary)

3.1 Modèle 1

Notre série est TS avec un unique changement structurel. Une série TS a la forme suivante:

$$Y_t = \mu + \delta tendance_t + u_t \quad (19)$$

La partie $\mu + \delta tendance_t$ est la partie déterministe et u_t est la partie aléatoire. u_t est un processus ARMA (Autoregressive Moving Average) stationnaire.

- $E(Y_t) = \mu + \delta tendance_t$: la moyenne dépend du temps
- $Var(Y_t) = \sigma_u^2$: la variance ne dépend pas du temps

On peut transformer une série issue d'un processus TS de telle sorte que la série transformée soit issue d'un PGD stationnaire en enlevant de la série originale la partie déterministe estimée par les MCO.

$$Y_t^* = Y_t - \widehat{\sigma} tendance_t - \widehat{\mu} \tag{20}$$

On estime Y_t^* avec le code suivant **Code:**

Tts<-length(ict); trend=1:Tts

reg<-lm(ict trend); summary(reg)

TSstar<-ict-regcoef$[1] - trend * regcoef$*[2]*

Ensuite on regarde l'ACF et le PACF de la série TSstar pour voir s'il y a de l'autocorrélation.

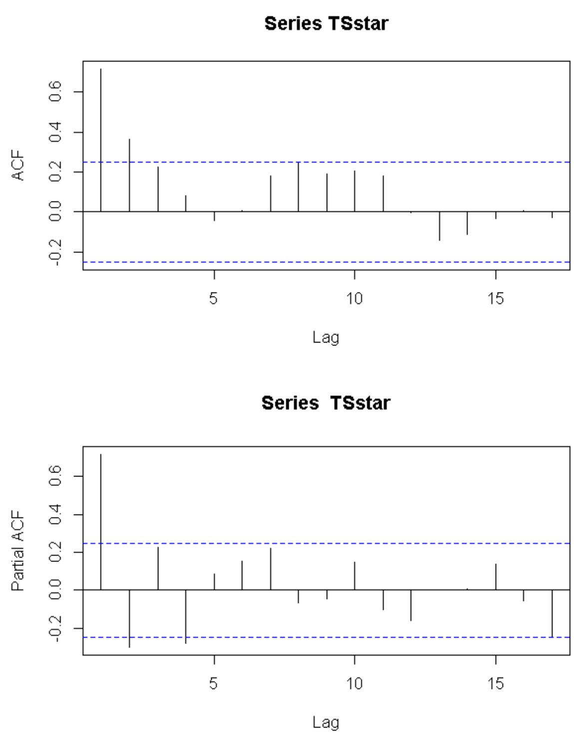

Figure 15: ACF, PACF

Selon l'ACF et le PACF, on de l'autocorrélation. On peut aussi tester s'il y a de l'autocorrélation avec le test de Ljung-Box.

```
> Box.test (TSstar, lag = 12, type = "Ljung-Box")

        Box-Ljung test

data:  TSstar
X-squared = 61.284, df = 12, p-value = 1.316e-08
```

Figure 16: Ljung-Box test

Comme la p-value est inférieure à 0.05, on rejette H_0 qui stipule qu'il n'y a pas de l'autocorrélation.

Comme la série est auto-corrélée, on utilise l'EACF pour déterminer l'ordre de p et q. L'EACF nous donne le résultat suivant:

```
> eacf(TSstar)
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 x x o o o o o o o o o  o  o  o
1 x o o o o o o o o o o  o  o  o
2 x o x o o o o o o o o  o  o  o
3 x o o o o o o o o o x  o  o  o
4 x x o o o o o o o o x  o  o  o
5 x o o o o o o o o o o  o  o  o
6 x o x o o o o o o o o  o  o  o
7 o o x o o x o o o o o  o  o  o
```

Figure 17: EACF

On trouve un ARMA(0,2) qui est de la forme suivante:

$$Y_t = \theta_0 + \epsilon_t - \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} \tag{21}$$

Il faut à présent tester si tous les coéfficients sont significatifs (p-value $<$ 0.05).

```
> reg<-Arima(TSstar,order=c(0,0,2))
ff qui sont significatifs
> coeftest(reg)

z test of coefficients:

           Estimate Std. Error z value  Pr(>|z|)    
ma1        1.088797   0.119722  9.0944 < 2.2e-16 ***
ma2        0.340630   0.110707  3.0768  0.002092 ** 
intercept -0.037442   0.462571 -0.0809  0.935487    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 18: ARMA(2,2) avec β_0

Comme la p-value de l'intercept est inférieure à 0.05, on l'enlève et on obtient le résultat suivant.

```
> reg<-Arima(TSstar,order=c(0,0,2),include.mean = FALSE)
aut avoir que des coeff qui sont significatifs
> coeftest(reg)

z test of coefficients:

    Estimate Std. Error z value  Pr(>|z|)
ma1  1.08880    0.11972  9.0946 < 2.2e-16 ***
ma2  0.34060    0.11071  3.0765  0.002095 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 19: ARMA(2,2) sans β_0

3.2 Tests d'autocorrélation sur les résidus

Maintenant que nous avons ce modèle avec les coéfficients θ_1 et θ_2 significatifs, on passe aux tests sur les résidus. On vérifie que l'aléa du modèle est bien un Bruit Blanc. Pour cela on va effectuer les trois tests suivants:

- T test
- ARCH test
- Ljung Box test

```
> t.test(reg$residuals)
on les résidus sont pas nuls

        One Sample t-test

data:  reg$residuals
t = -0.015288, df = 61, p-value = 0.9879
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.3912931  0.3853551
sample estimates:
   mean of x
-0.002968999
```

Figure 20: T test sur résidus

Comme la p-value (0.9879) est supérieure à 5%, on ne rejette pas H_0 qui stipule que l'espérance des aléas est nulle.

```
> ArchTest(reg$residuals, lag=12)

        ARCH LM-test; Null hypothesis: no ARCH effects

data:  reg$residuals
Chi-squared = 11.152, df = 12, p-value = 0.516
```

Figure 21: ARCH test sur résidus

Comme la p-value (0.516) est supérieure à 5%, on ne rejette pas H_0 qui stipule qu'il n'y a pas de cluster de volatilité.

```
> Box.test (reg$residuals, lag = 20, type = "Ljung")

        Box-Ljung test

data:  reg$residuals
X-squared = 18.21, df = 20, p-value = 0.5736
```

Figure 22: Ljung Box test sur résidus

Comme la p-value (0.5736) est supérieure à 5%, on ne rejette pas H_0 qui stipule qu'il n'y a pas d'autocorrélation jusqu'à l'ordre 20. On conclut que l'aléa est bien un Bruit Blanc.

Le BIC obtenu pour ce modèle est 241.3081.

3.3 Modèle 2

On essaye d'autres valeurs de p et q d'ARMA afin de minimiser le BIC. On obtient un ARMA(1,1) qui est de la forme suivante:

$$Y_t - \phi_1 Y_{t-1} = \phi_0 + \epsilon_t - \theta_1 \epsilon_{t-1} \tag{22}$$

avec $\phi_1 \neq \theta_1$

```
> reg<-Arima(TSstar,order=c(1,0,1))
> coeftest(reg)

z test of coefficients:

           Estimate Std. Error z value  Pr(>|z|)
ar1        0.454449   0.129521  3.5087 0.0004503 ***
ma1        0.652187   0.108876  5.9902 2.096e-09 ***
intercept -0.065317   0.564980 -0.1156 0.9079621
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 23: ARMA(1,1) avec β_0

Comme la p-value de l'intercept est inférieure à 0.05, on l'enlève et on obtient le résultat suivant.

```
> reg<-Arima(TSstar,order=c(1,0,1),include.mean = FALSE)
> coeftest(reg)

z test of coefficients:

    Estimate Std. Error z value  Pr(>|z|)
ar1  0.45418    0.12952  3.5067 0.0004536 ***
ma1  0.65239    0.10880  5.9963 2.019e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 24: ARMA(1,1) sans β_0

3.4 Tests d'autocorrélation sur les résidus

```
> t.test(reg$residuals)

        One Sample t-test

data:  reg$residuals
t = -0.019704, df = 61, p-value = 0.9843
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.3869609  0.3794094
sample estimates:
   mean of x
-0.003775788
```

Figure 25: T test sur résidus

Comme la p-value (0.9843) est supérieure à 5%, on ne rejette pas H_0 qui stipule que l'espérance des aléas est nulle.

```
> ArchTest(reg$residuals, lag=12)

        ARCH LM-test; Null hypothesis: no ARCH effects

data:  reg$residuals
Chi-squared = 10.438, df = 12, p-value = 0.5776
```

Figure 26: ARCH test sur résidus

Comme la p-value (0.5776) est supérieure à 5%, on ne rejette pas H_0 qui stipule qu'il n'y a pas de cluster de volatilité.

```
> Box.test (reg$residuals, lag = 20, type = "Ljung")

        Box-Ljung test

data:  reg$residuals
X-squared = 18.9, df = 20, p-value = 0.5284
```

Figure 27: Ljung Box test sur résidus

Comme la p-value (0.5284) est supérieure à 5%, on ne rejette pas H_0 qui stipule qu'il n'y a pas d'autocorrélation jusqu'à l'ordre 20. On conclut que l'aléa est bien un Bruit Blanc.

Le BIC obtenu pour ce modèle est 239.6366.

On retient le modèle ARMA(1,1) car son BIC est plus petit.

4 Prévision

En retenant le modèle ARMA(1,1), nous avons les prévisions suivantes:

```
> fc <- forecast(TSstar, h=5,model=reg)
> fc
     Point Forecast     Lo 80     Hi 80     Lo 95    Hi 95
2022    -1.08846016 -3.038229 0.8613082 -4.070374 1.893453
2023    -0.49435720 -3.402385 2.4136705 -4.941802 3.953087
2024    -0.22452732 -3.293218 2.8441635 -4.917685 4.468630
2025    -0.10197589 -3.202773 2.9988209 -4.844235 4.640284
2026    -0.04631545 -3.153694 3.0610629 -4.798641 4.706010
```

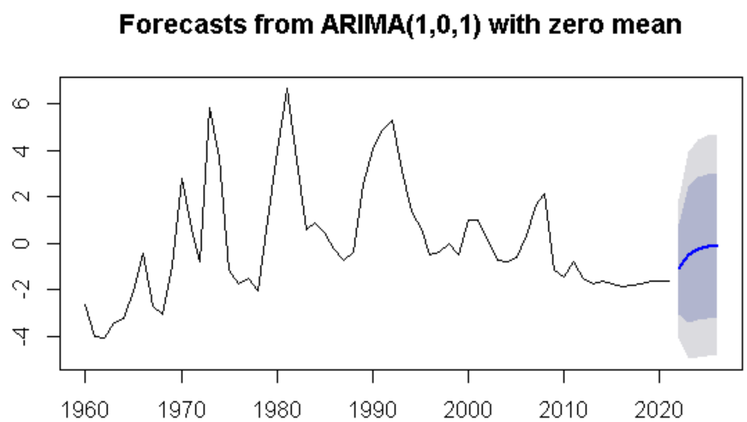

Figure 28: Prévisions de 2022 à 2026

Le taux d'intérêt de court terme prévu pour 2022 est de -1.088 %. Cependant comme l'intervalle de confiance est assez grande, les prévisions ne sont pas très précises.