



Économétrie Appliquée : Recueil des cas pratiques sur EViews et Stata

Jonas Kibala Kuma

► To cite this version:

Jonas Kibala Kuma. Économétrie Appliquée : Recueil des cas pratiques sur EViews et Stata. Licence. Congo-Kinshasa. 2018. cel-01771070

HAL Id: cel-01771070

<https://hal.science/cel-01771070>

Submitted on 19 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Kinshasa, Avril 2018
(1^{ère} édition)

Manuel d'Econométrie
(Inspiré de Fodiye Bakary Doucoure, 2008)

Économétrie Appliquée :
Recueil des cas pratiques sur EViews et Stata

Par

Jonas KIBALA KUMA

(Licencié en Sciences Economiques, DEA-PTC/Unikin en cours)

Centre de Recherches Economiques et Quantitatives
(CREQ)

« *Editions Ecodata* »

Kinshasa, Avril 2018

Manuel d'Econométrie
(Inspiré de Fodiye Bakary Doucoure, 2008)

Économétrie Appliquée : Recueil des cas pratiques sur EViews et Stata

Par

Jonas KIBALA KUMA

(DEA-PTC Economie/Unikin en cours)

–

*Centre de Recherches Economiques et Quantitatives
(CREQ)*

– 1^{ère} édition –

« **Editions Ecodata** »



– Note –

Ce manuel d'Econométrie, qui s'inscrit dans le cadre de nos travaux ou recherches sur « l'Econométrie appliquée » que nous nous efforçons de mettre à la disposition des chercheurs africains (congolais surtout) de tout bord, est un recueil des travaux pratiques qui s'inspire largement de l'ouvrage du Professeur FODIYE BAKARY Doucoure (2008) intitulé « Méthodes économétriques : cours et travaux pratiques ». Ce manuel, riche en illustrations et à compter parmi les rares des éditions africaines qui combinent la théorie et la pratique sur logiciels, est à mon avis indispensable lorsque l'on se propose d'apprendre l'Econométrie en théorie et en pratique à l'aide des logiciels « Eviews et Stata ».

Dans ce manuel, nous ne nous limitons pas seulement à résumer les aspects théoriques tant soit peu, mais aussi, sur base de mêmes séries et variables (soit nos propres séries), nous faisons des analyses particulières de nature à s'ajouter ou compléter celles du Prof.

Par ailleurs, signalons qu'en plus de l'ouvrage du Professeur FODIYE BAKARY Doucoure, qui constitue notre pilori, nous avons consulté bien d'autres documents (Veuillez consulter nos références bibliographiques) pour la plupart axés sur la pratique de techniques économétriques et statistiques sur EViews et/ou Stata. Par ce manuel, nous espérons vous familiariser à la pratique de l'économétrie sur ces deux logiciels.

_____Merci au Professeur FODIYE BAKARY Doucoure (Université Cheick Anta Diop du Sénégal) pour ses travaux de recherche qui nous inspirent.

Kinshasa, Avril 2018



« Il est à peu près impossible de faire de la recherche en sciences économiques sans se trouver devant la nécessité de lire ou de réaliser des travaux d'économétrie à un moment ou un autre » (Fodiye B.D., p.3).



CHAP I MODELE LINEAIRE GENERAL

Cas pratique 1.1 : importation fonction du PIB

► **Modèle :**

$$\log(IMP_t) = a_0 + a_1 \log(PIBR_t) + e_t \dots \dots [1.1]$$

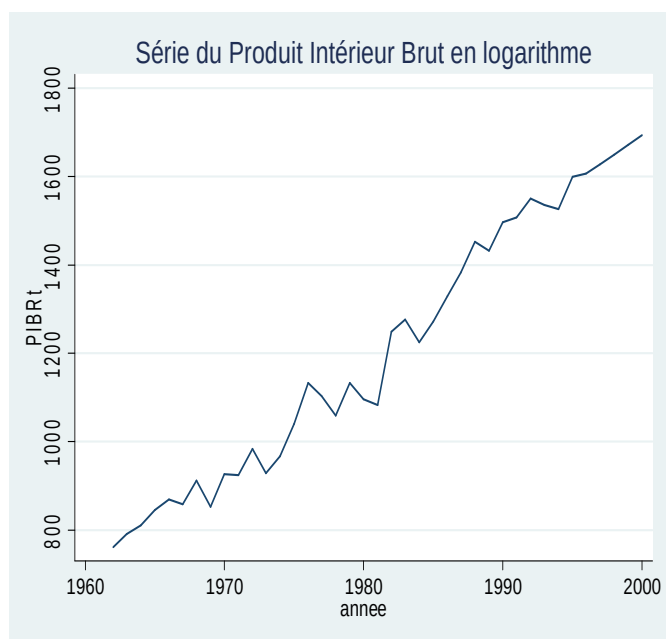
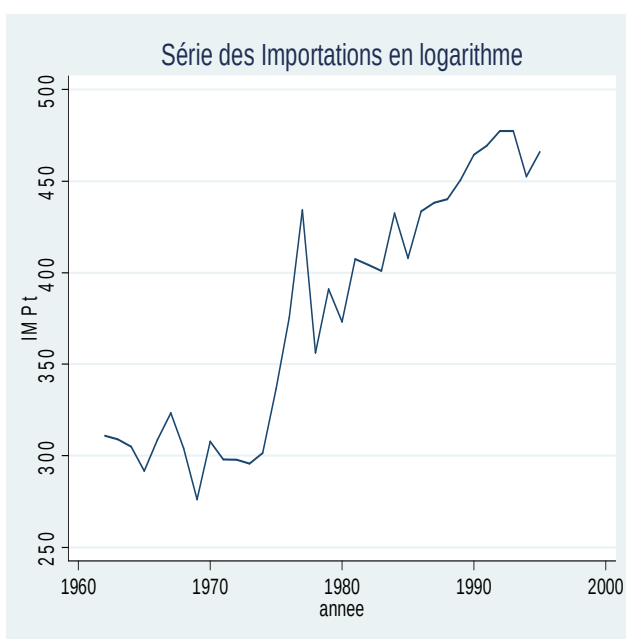
► **Travail demandé (recours à Stata et Eviews) :**

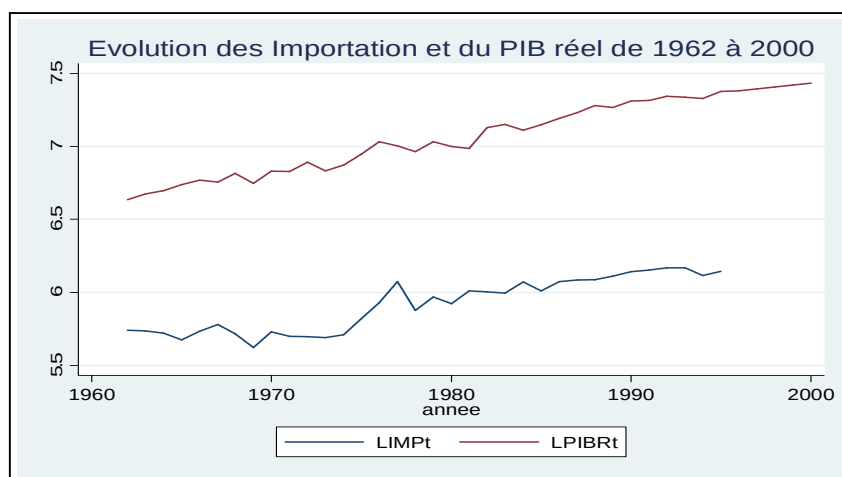
- o Représenter graphiquement les variables IMP_t et $PIBR_t$, en obtenir les caractéristiques et Tester leur normalité/log-normalité (test de Jarque Bera) ;
- o Estimer les coefficients/paramètres « a_0 et a_1 » par les MCO et interpréter les économiquement (interpréter aussi le R^2) ;
- o Effectuer les tests usuels (significativité individuelle et globale des paramètres, hétéroscédasticité, autocorrélation des erreurs, bonne spécification et stabilité des coefficients) et faire de corrections éventuelles (corriger l'autocorrélation et stabiliser le modèle) ;
- o Prévoir les importations pour les années 1996, 1997, 1998, 1999 et 2000.

a) Représentation graphique des variables IMP_t et $PIBR_t$ et Teste de leur normalité et log-normalité (test de Jarque Bera)

► **Les commandes Stata pour les graphiques (line = tsline) :**

```
gen LIMPt=log(IMPt)
gen LPIBRt=log(PIBRt)
line IMPt annee, title(Série des Importations en logarithme)
line PIBRt annee, title(Série du Produit Intérieur Brut en logarithme)
twoway (line LIMPt LPIBRt annee),title(Evolution des Importation et du PIB
réel de 1962 à 2000)
```





► **Normalité et log-normalité de variables (test de Jarque-Berra/Shapiro wilk) :**

_____ Sur Eviews, le chemin est (test de Jarque-Berra) : Quick/Group Statistics/Descriptive Statistics/Common sample → dans series list, taper : pibr lpibr imp limp → ok :

	PIBR	LPIBR	IMP	LIMP
Mean	1144.376	7.017415	377.0147	5.916234
Median	1099.200	7.002333	383.2000	5.948355
Maximum	1599.400	7.377384	477.3000	6.168145
Minimum	761.3000	6.635027	276.2000	5.621125
Std. Dev.	260.6200	0.228128	67.88812	0.182731
Skewness	0.274254	0.043221	0.026034	-0.089097
Kurtosis	1.740300	1.709849	1.443741	1.426565
Jarque-Bera	2.674250	2.368611	3.434927	3.552221
Probability	0.262600	0.305959	0.179521	0.169295
Sum	38908.80	238.5921	12818.50	201.1520
Sum Sq. Dev.	2241451.	1.717400	152090.3	1.101893
Observations	34	34	34	34

H_0 : Normalité de la variable (prob>5%)

H_1 : Non normalité de la variable (prob<5%)

Commentaire : toutes les variables sont normalement distribuées (elles suivent les lois normale et *log-normale*). Ex : il y a 26,26% de chance de prendre une mauvaise décision en optant pour la non normalité de la variable « PIB ».

_____ Sur Stata, taper (test de Shapiro-Wilk, 1995) :

sktest PIBRt LPIBRt IMPt LIMPt

Skewness/Kurtosis tests for Normality				
Variable	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	joint Prob>chi2
PIBRt	0.666	0.000	12.98	0.0015
LPIBRt	0.820	0.000	11.74	0.0028
IMPt	0.943	0.000	18.61	0.0001
LIMPt	0.807	0.000	19.84	0.0000

H_0 : la variable est gaussienne ($W_c > W_t$)

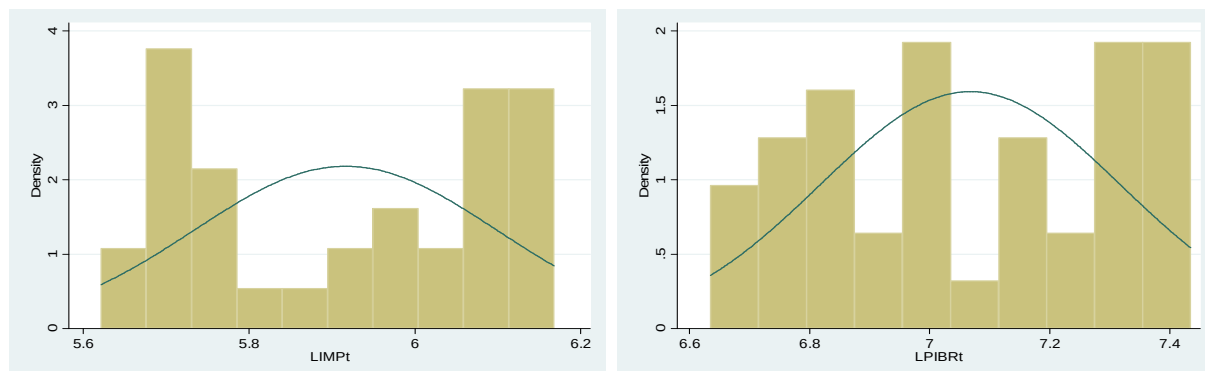
H_1 : la variable est non gaussienne ($W_c < W_t$)



Les histogrammes respectifs de variables peuvent nous aider aussi : sur Stata, taper :

`hist LIMPt, normal bin(10)`

`hist LPIBRt, normal bin(10)`



► Caractéristiques/statistiques descriptives de variables

Sur Eviews, Cfr le chemin relatif au test de Jarque-Berra ci-dessus ;

Sur stata, la commande est : `sum IMPt LIMPt PIBRt LPIBRt, detail`

(NB : ci-dessous, nous ne présentons que le résultat relatif à la variable « IMPt », pour raison d'espace et à titre illustratif)

IMPt					

	Percentiles	Smallest			
1%	276.2	276.2			
5%	291.6	291.6			
10%	297.8	295.7	Obs		34
25%	307.8	297.8	Sum of Wgt.		34
50%	383.2		Mean		377.0147
		Largest	Std. Dev.		67.88812
75%	438.2	466.2	Variance		4608.797
90%	466.2	469.2	Skewness		.026034
95%	477.2	477.2	Kurtosis		1.443741
99%	477.3	477.3			

b) Estimation des coefficients/paramètres « a_0 et a_1 » par les MCO (sur Stata) et interprétation économique

Commande: `reg LIMPt LPIBRt`

Source	SS	df	MS			
Model	.967660088	1	.967660088	Number of obs =	34	
Residual	.134233167	32	.004194786	F(1, 32) =	230.68	
				Prob > F	= 0.0000	
				R-squared	= 0.8782	
				Adj R-squared	= 0.8744	
Total	1.10189325	33	.033390705	Root MSE	= .06477	
LIMPt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
LPIBRt	.7506297	.0494219	15.19	0.000	.6499607	.8512988
_cons	.6487534	.3469917	1.87	0.071	-.0580455	1.355552

Commentaires :

- o Paramètre : une variation du Produit Intérieur Brut Réel/PIBR, dans le sens positif, de 10% entraîne une modification des importations de 7,5% dans le



même sens (a_1 = élasticité ou paramètre de sensibilité relative des importations par rapport au mouvement du PIBR);

- o Coefficient de détermination: le modèle --- tel que spécifié (variable explicative/importation, relation linéaire, etc.) --- explique 87,82% de variations/évolutions des importations ($R^2 = 0.8782$);
- o Coefficient de corrélation (simple): nos deux variables (LIMPt et LPIBRt) sont fortement et positivement corrélées à 93,71% (cette relation est linéaire), car : $prob < 5\%$.

_____ Les hypothèses du test:

H_0 : Pas de corrélation linéaire ($prob-F > 5\%$)

H_1 : Existence d'une corrélation linéaire ($prob-F < 5\%$)

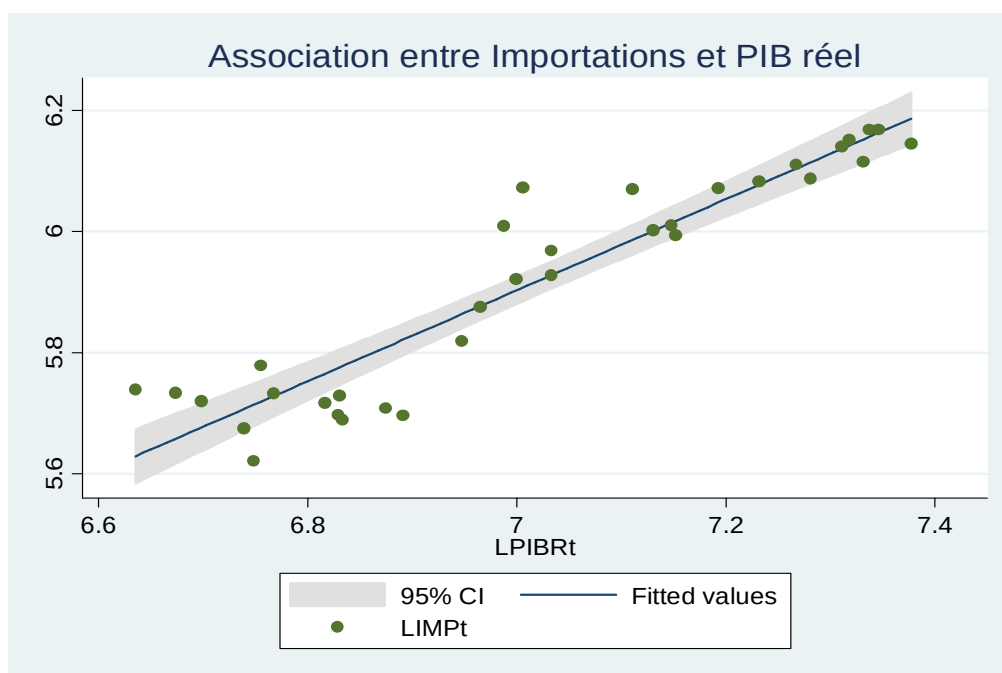
Commande: `pwcorr LIMPt LPIBRt, sig print(1)`

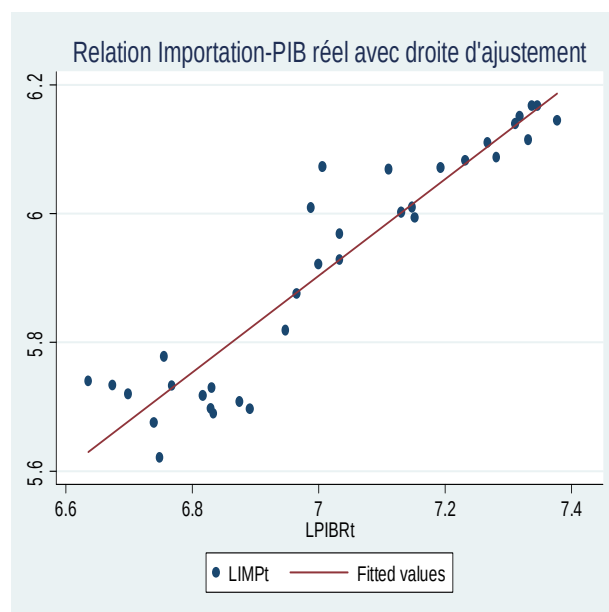
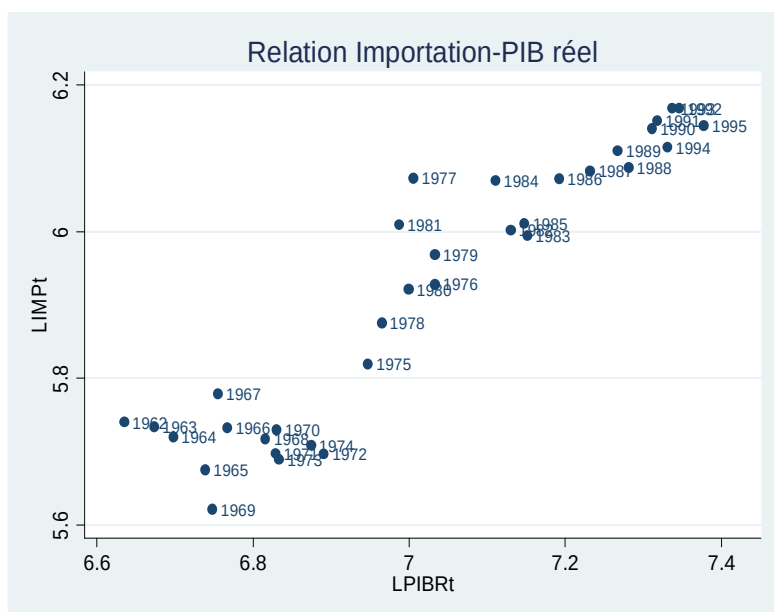
	LIMPt	LPIBRt
LIMPt	1.0000	
LPIBRt	0.9371 0.0000	1.0000

_____ Corrélation de variables sur le graphique:

Commandes Stata:

- o `twoway (lfitci LIMPt LPIBRt) (scatter LIMPt LPIBRt), title(Association entre Importations et PIB réel)`
- o `twoway scatter LIMPt LPIBRt, mlabel(annee) title(Relation Importation-PIB réel)`
- o `scatter LIMPt LPIBRt || lfit LIMPt LPIBRt, title(Relation Importation-PIB réel avec droite d'ajustement)`





c) Tests usuels et corrections éventuelles

- ✚ **Significativité individuelle** : Les paramètres estimés « a_0 et a_1 » sont statistiquement significatifs aux seuils de 10% et 1% (au regard des probabilités statistiques associées), respectivement ;
- ✚ **Significativité conjointe** : Pris ensemble, les paramètres estimés « a_0 et a_1 » sont statistiquement significatifs au regard de la statistique calculée de Fisher ($\text{prob } F\text{-stat} < 5\%$) ;
- ✚ **Normalité des erreurs** : Les erreurs sont normalement distribuées ($\text{prob} > 5\%$).

Les hypothèses du test :

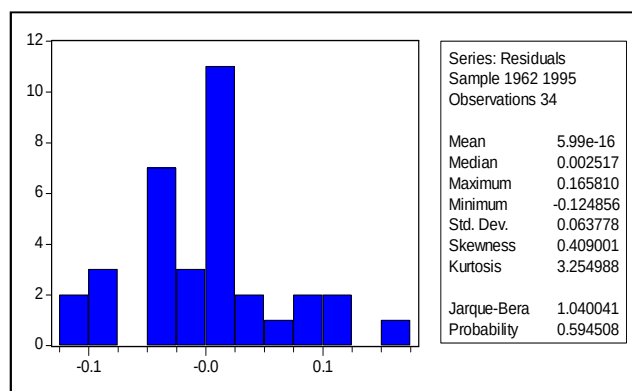
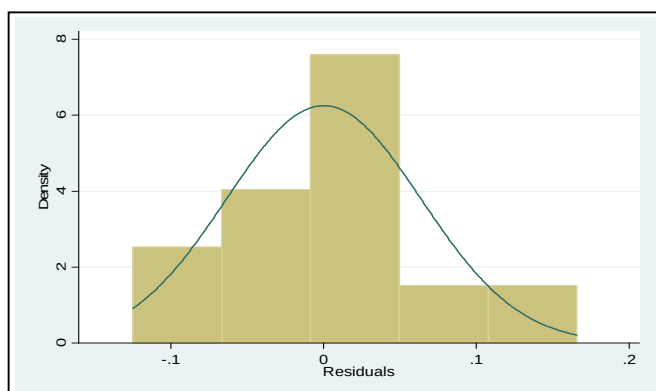
- $\left\{ \begin{array}{l} H_0 : \text{Les erreurs sont normalement distribuées (prob-F} > 5\%) \\ H_1 : \text{Les erreurs ne sont pas normalement distribuées (prob-F} < 5\%) \end{array} \right.$

Commandes Stata (pour l'histogramme) :

- $\left\{ \begin{array}{l} \text{reg LIMPt LPIBt} \\ \text{predict e, resid} \\ \text{hist e, normal} \end{array} \right.$

Commande Eviews :

Dans l'output, suivre le chemin : View/Residual Tests/Histogram-Normality Test.



Commande: `sktest e`

Skewness/Kurtosis tests for Normality				
Variable	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	Prob>chi2
e	0.271	0.410	2.03	0.3619

✚ Hétéroscédasticité :

(i) **Test de white** : Les erreurs sont homoscedastiques, car : prob-F > 5% et tous les paramètres estimés sont statistiquement non significatifs.

Les hypothèses du test :

$\begin{cases} H_0 : \text{Homoscedasticité (prob-F > 5\%)} \\ H_1 : \text{Absence d'homoscedasticité (prob-F < 5\%)} \end{cases}$

Commandes Stata :

$\begin{cases} \text{predict res, resid} \\ \text{gen res2=res^2} \\ \text{gen LPIBrt2=LPIBrt^2} \\ \text{reg res2 LPIBrt LPIBrt2} \end{cases}$

Source	SS	df	MS	Number of obs = 34		
Model	.000185542	2	.000092771	F(2, 31)	=	2.85
Residual	.001009501	31	.000032565	Prob > F	=	0.0731
Total	.001195043	33	.000036213	R-squared	=	0.1553
				Adj R-squared	=	0.1008
				Root MSE	=	.00571
res2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
LPIBrt	.3005092	.323429	0.93	0.360	-.3591285	.960147
LPIBrt2	-.0220746	.0230267	-0.96	0.345	-.0690379	.0248887
_cons	-1.016689	1.134748	-0.90	0.377	-3.331024	1.297646

(ii) **Test de présence des effets ARCH d'ordre 3 (ARCH(3) ou LM test)** : Les erreurs sont homoscedastiques, car : prob-F > 5% et tous les paramètres estimés sont statistiquement non significatifs.

Les hypothèses du test :

H_0 : Homoscedasticité ou absence d'effets ARCH (prob-F > 5%)

H_1 : Absence d'homoscedasticité ou présence d'effets ARCH (prob-F < 5%)

Commande: `reg res2 L.res2 L2.res2 L3.res2`

Source	SS	df	MS	Number of obs = 31		
Model	.000032965	3	.000010988	F(3, 27)	=	0.27
Residual	.001084147	27	.000040154	Prob > F	=	0.8438
Total	.001117112	30	.000037237	R-squared	=	0.0295
				Adj R-squared	=	-0.0783
				Root MSE	=	.00634
res2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
res2						
L1.	-.031465	.1898109	-0.17	0.870	-.4209248	.3579949
L2.	.045063	.1900358	0.24	0.814	-.3448581	.4349842
L3.	.160971	.1856564	0.87	0.394	-.2199644	.5419064
_cons	.0029522	.0017273	1.71	0.099	-.000592	.0064963



- (iii) **Test de Breusch-Pagan** : Les erreurs ne sont pas homoscedastiques (elles sont hétéroscedastiques), car : $\text{prob-}\chi^2(1) < 5\%$.

Les hypothèses du test :

H_0 : Homoscedasticité ($\text{prob-F} > 5\%$ et $\chi^2_{dl(c)} < \chi^2_{dl(T)}$)

H_1 : Absence d'homoscedasticité ($\text{prob-F} < 5\%$ et $\chi^2_{dl(c)} > \chi^2_{dl(T)}$)

Commande : estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

H_0 : Constant variance

Variables: fitted values of LIMPt

$\chi^2(1) = 4.99$

Prob > χ^2 = 0.0255

► **Autocorrélation des erreurs :**

- (i) **Test de Durbin-Watson** : les erreurs semblent positivement auto-corrélées à l'ordre 1 (présomption), car $\text{prob} < 5\%$.

Les hypothèses du test :

H_0 : Absence d'autocorrélation ($\text{prob} > 5\% \leftrightarrow \rho = 0$)

H_1 : Présence d'autocorrélation ($\text{prob} < 5\% \leftrightarrow \rho \neq 0$)

NB : $n=34$ et $k=1$ (nombre de régresseurs) ; au seuil de 5%, sur la table de Durbin-Watson : $d_{\text{inf}} = 1,39$ et $d_{\text{sup}} = 1,51$. En outre, $DW = 1,064 < d_{\text{inf}}$.

o Calcul de la statistique de Durbin-Watson :

Commande:dwstat ou estat dwatson

Durbin-Watson d-statistic(2, 34) = 1.064825

o Test de Durbin-Watson :

Commande:durбина, lags(1)

Durbin's alternative test for autocorrelation

lags(p)	chi2	df	Prob > chi2
1	6.592	1	0.0102
H0: no serial correlation			

- (ii) **Test de Breush-Godfrey** : les erreurs sont auto-corrélées à l'ordre 2, car $\text{prob} < 5\%$.

Les hypothèses du test :

H_0 : Absence d'autocorrélation ($\text{prob} > 5\%$)

H_1 : Présence d'autocorrélation ($\text{prob} < 5\%$)

Commande:estat bgodfrey, lags(3)

Breusch-Godfrey LM test for autocorrelation

lags(p)	chi2	df	Prob > chi2
3	6.451	3	0.0916



Commande: estat bgodfrey, lags(2)

Breusch-Godfrey LM test for autocorrelation

lags(p)	chi2	df	Prob > chi2
2	6.380	2	0.0412
H0: no serial correlation			

ou encore

Commande: bgodfrey, lags(3) small

Breusch-Godfrey LM test for autocorrelation

lags(p)	F	df	Prob > F
3	2.150	(3, 29)	0.1155
H0: no serial correlation			

Commande: bgodfrey, lags(2) small

Breusch-Godfrey LM test for autocorrelation

lags(p)	F	df	Prob > F
2	3.190	(2, 30)	0.0555
H0: no serial correlation			

(iii) **Test de présence d'effets ARCH(2)** : Il y a absence d'effets ARCH (à l'ordre 1, 2, ...) dans le modèle estimé/résidus (no ARCH effects), car $prob > 5\%$.

____ Les hypothèses du test :

H_0 : Les erreurs ne suivent pas un modèle ARCH(1), ARCH(2) $\rightarrow (prob > 5\%)$

H_1 : Les erreurs suivent un modèle ARCH(1), ARCH(2) $\rightarrow (prob < 5\%)$

Commande: archlm, lags(1)

LM test for autoregressive conditional heteroskedasticity (ARCH)

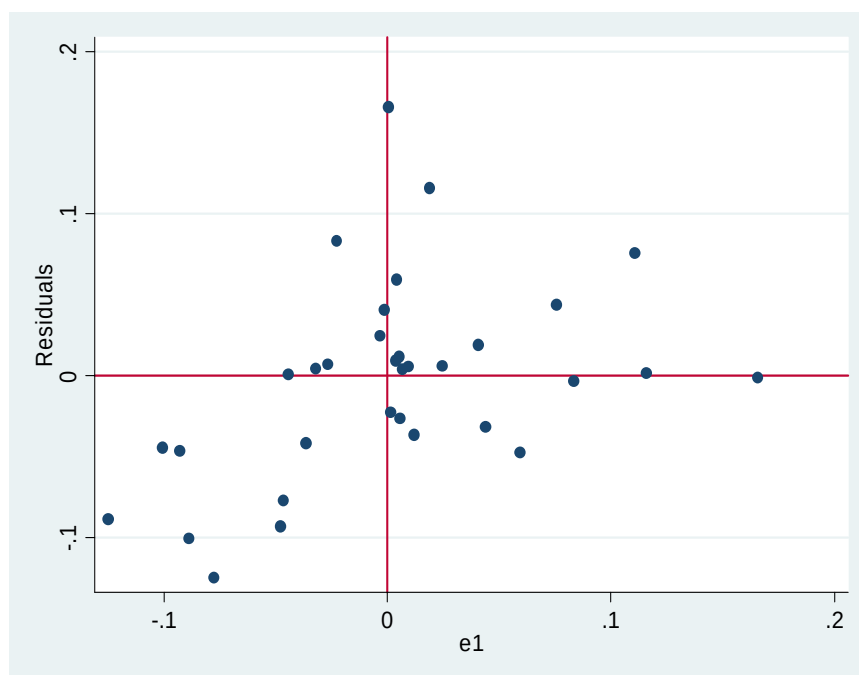
lags(p)	chi2	df	Prob > chi2
1	0.009	1	0.9230
H0: no ARCH effects vs. H1: ARCH(p) disturbance			

Commande: archlm, lags(2)

LM test for autoregressive conditional heteroskedasticity (ARCH)

lags(p)	chi2	df	Prob > chi2
2	0.057	2	0.9719
H0: no ARCH effects vs. H1: ARCH(p) disturbance			



(iv) Evolution des résidus dans le temps (par rapport à ses valeurs passées) :**Commande:** `graph twoway (scatter e e1), yline(0) xline(0)`Ou simplement : `scatter e e1, yline(0) xline(0)`**(v) Correction de l'autocorrélation des erreurs**✚ **Approche de Cochrane-Orcutt (1949) :****Commande (Stata):** `prais LIMPt LPIBRt, corc`

```

Iteration 0: rho = 0.0000
Iteration 1: rho = 0.4210
Iteration 2: rho = 0.4143
Iteration 3: rho = 0.4141
Iteration 4: rho = 0.4141
Iteration 5: rho = 0.4141
Cochrane-Orcutt AR(1) regression -- iterated estimates

```

Source	SS	df	MS			
Model	.337035999	1	.337035999	Number of obs =	33	
Residual	.097710835	31	.003151962	F(1, 31) =	106.93	
Total	.434746833	32	.013585839	Prob > F =	0.0000	
				R-squared	= 0.7752	
				Adj R-squared	= 0.7680	
				Root MSE	= .05614	

LIMPt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
LPIBRt	.7745266	.0749012	10.34	0.000	.6217646	.9272886
_cons	.4737872	.5279349	0.90	0.376	-.6029432	1.550518
rho	.4140699					


```

Durbin-Watson statistic (original)    1.064825
Durbin-Watson statistic (transformed) 2.181396

```



Sur Eviews, faire : `ls limp c lpibr ar(1)`

Dependent Variable: LIMP				
Method: Least Squares				
Date: 11/30/13 Time: 15:41				
Sample (adjusted): 1963 1995				
Included observations: 33 after adjustments				
Convergence achieved after 6 iterations				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.473789	0.536888	0.882473	0.3845
LPIBR	0.774526	0.076162	10.16940	0.0000
AR(1)	0.414069	0.156776	2.641142	0.0130
R-squared	0.908666	Mean dependent var	5.921581	
Adjusted R-squared	0.902577	S.D. dependent var	0.182844	
S.E. of regression	0.057070	Akaike info criterion	-2.802554	
Sum squared resid	0.097711	Schwarz criterion	-2.666508	
Log likelihood	49.24214	F-statistic	149.2321	
Durbin-Watson stat	2.181395	Prob(F-statistic)	0.000000	
Inverted AR Roots	.41			

 **Méthode de Prais-Winstern :**

Commande: `prais LIMPt LPIBRt, sseseach`

```
Iteration 1: rho = 0.8944 , criterion = -.12478055
Iteration 2: rho = 0.4894 , criterion = -.1072125
Iteration 3: rho = 0.4894 , criterion = -.1072125
Iteration 4: rho = 0.4894 , criterion = -.1072125
Iteration 5: rho = 0.4894 , criterion = -.1072125
Iteration 6: rho = 0.4891 , criterion = -.10721247
Iteration 7: rho = 0.4886 , criterion = -.10721244
Iteration 8: rho = 0.4886 , criterion = -.10721244
Iteration 9: rho = 0.4886 , criterion = -.10721244
Iteration 10: rho = 0.4886 , criterion = -.10721244
Iteration 11: rho = 0.4886 , criterion = -.10721244
Iteration 12: rho = 0.4886 , criterion = -.10721244
Iteration 13: rho = 0.4886 , criterion = -.10721244
Iteration 14: rho = 0.4886 , criterion = -.10721244
```

Prais-Winsten AR(1) regression -- SSE search estimates

Source	SS	df	MS	Number of obs = 34		
Model	4.03283805	1	4.03283805	F(1, 32)	=	1203.69
Residual	.107212439	32	.003350389	Prob > F	=	0.0000
Total	4.14005049	33	.125456075	R-squared	=	0.9741
				Adj R-squared	=	0.9733
				Root MSE	=	.05788
LIMPt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
LPIBRt	.7044378	.0773969	9.10	0.000	.5467855	.8620901
_cons	.9747038	.5434083	1.79	0.082	-.1321826	2.08159
rho	.4886445					
Durbin-Watson statistic (original) 1.064825						
Durbin-Watson statistic (transformed) 2.148891						



Approche de White :

Sur Eviews, dans l'output, suivre : Estimate/Options→Cocher « Heteroskedasticity consistent coefficient covariance »; ensuite, cocher « White »→ok :

Dependent Variable: LIMP				
Method: Least Squares				
Date: 11/30/13 Time: 16:04				
Sample (adjusted): 1962 1995				
Included observations: 34 after adjustments				
White Heteroskedasticity-Consistent Standard Errors & Covariance				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.648755	0.312587	2.075437	0.0461
LPIBR	0.750630	0.043554	17.23457	0.0000
R-squared	0.878179	Mean dependent var	5.916234	
Adjusted R-squared	0.874373	S.D. dependent var	0.182731	
S.E. of regression	0.064767	Akaike info criterion	-2.579013	
Sum squared resid	0.134233	Schwarz criterion	-2.489227	
Log likelihood	45.84322	F-statistic	230.6815	
Durbin-Watson stat	1.064827	Prob(F-statistic)	0.000000	

Sur Stata, taper (correction de l'hétéroscédasticité par la méthode d'Eicker-White ou la méthode des écarts robustes): reg LIMPt LPIBRt, robust

Linear regression					Number of obs =	34
					F(1, 32) =	297.03
					Prob > F =	0.0000
					R-squared =	0.8782
					Root MSE =	.06477

		Robust				
LIMPt		Coef.	Std. Err.	t	P> t	[95% Conf. Interval]

LPIBRt		.7506297	.0435538	17.23	0.000	.6619136 .8393459
_cons		.6487534	.3125875	2.08	0.046	.0120336 1.285473

Approche de Newey-West :

Sur Eviews, dans l'output, suivre : Estimate/Options→Cocher « Heteroskedasticity consistent coefficient covariance », ensuite, cocher « Newey-West »→ok :

Dependent Variable: LIMP				
Method: Least Squares				
Date: 11/30/13 Time: 16:16				
Sample (adjusted): 1962 1995				
Included observations: 34 after adjustments				
Newey-West HAC Standard Errors & Covariance (lag truncation=3)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.648755	0.442118	1.467379	0.1520
LPIBR	0.750630	0.061301	12.24503	0.0000
R-squared	0.878179	Mean dependent var	5.916234	
Adjusted R-squared	0.874373	S.D. dependent var	0.182731	
S.E. of regression	0.064767	Akaike info criterion	-2.579013	
Sum squared resid	0.134233	Schwarz criterion	-2.489227	
Log likelihood	45.84322	F-statistic	230.6815	
Durbin-Watson stat	1.064827	Prob(F-statistic)	0.000000	



Sur Stata, Commandes :

newey LIMPt LPIBRt, lag(1)

newey LIMPt LPIBRt, lag(2)

Regression with Newey-West standard errors				Number of obs =	34
maximum lag: 1				F(1, 32) =	197.54
				Prob > F =	0.0000

LIMPt	Coef.	Newey-West Std. Err.	t	P> t	[95% Conf. Interval]

LPIBRt	.7506297	.0534069	14.05	0.000	.6418435 .859416
_cons	.6487534	.3841783	1.69	0.101	-.1337922 1.431299

Regression with Newey-West standard errors				Number of obs =	34
maximum lag: 2				F(1, 32) =	166.58
				Prob > F =	0.0000

LIMPt	Coef.	Newey-West Std. Err.	t	P> t	[95% Conf. Interval]

LPIBRt	.7506297	.058158	12.91	0.000	.6321657 .8690937
_cons	.6487534	.4190124	1.55	0.131	-.204747 1.502254

📊 Comparaison de résultats obtenus par les 4 approches

Méthodes/ Approches	Somme des carrés moyens des erreurs/RMSE	Coefficient (R ²)	Statistique de Fisher	Durbin-Watson
Cochrane-Orcut	0.05614	77.52%	106.93	2.181396
Prais-Winstern	0.05788	97.41%	1203.69	2.148891
White	0.134233	87.82%	230.68	1.064827
Newey-West	0.134233	87.82%	230/198.54	1.064822

Commentaire : l'on notera que la méthode de White et celle de Newey-West servent plutôt à corriger un modèle victime d'hétéroscédasticité : Qu'elles ne nous rendent pas service, nous ne leur tenons pas rigueur. L'arbitrage est donc à faire entre les deux autres approches (Cochrane et Prais). Comme on peut le constater, d'une part, l'approche de Prais-Winstern l'emporte sur celle de Cochrane-Orcut si l'on s'en tient à la statistique de Fisher et au coefficient de détermination. Toutefois, les propos sont à nuancer au regard du critère RMSE le plus minimum offert par la méthode de Cochrane-Orcut, ce qui la rend d'autre part efficace par rapport à la méthode de Prais-Winstern. Et Alors ? Quant à moi, optons pour la méthode de prais-winstern, elle qui a réussi au moins à maintenir les paramètres statistiquement significatifs (seuils : 1% et 10%).

En outre, précisons ce qui suit : $n=33$ et $k=2$ [nombre de régresseurs : y compris AR(1)] ; au seuil de 5%, sur la table de Durbin-Watson : $d_{inf} = 1,32$ et $d_{sup} = 1,58$. Or, $DW = 2,148891 \Leftrightarrow 4 - d_{inf} < DW < 4 - d_{sup}$, avec $4 - d_{inf} = 2,68$ et $4 - d_{sup} = 2,42$. D'où, il y a absence d'autocorrélation des erreurs (l'approche de Prais-Winstern nous a aidée à corriger ce biais).



► **Bonne spécification du modèle (test Reset de Ramsey) :**

Les hypothèses du test :

H_0 : Le modèle est bien spécifié $\rightarrow (prob > 5\%)$

H_1 : Le modèle n'est pas bien spécifié $\rightarrow (prob < 5\%)$

Commande : $\begin{cases} \text{reg LIMPt LPIBRt} \\ \text{ovtest} \end{cases}$

Ramsey RESET test using powers of the fitted values of LIMPt
 H_0 : model has no omitted variables
 $F(3, 29) = 6.57$
Prob > F = 0.0016

Autrement, estimer le modèle (1) ci-dessous par les MCO et tester la significativité de paramètres associés aux régresseurs (avec $LIMPF3_t = LIMPt$ prévu élevé au cube, etc.) :

$$LIMPt = a_0 + a_1LPIBR_t + a_2LIMPF2_t + a_3LIMPF3_t + e_t \dots [1]$$

Commandes Stata :

$\begin{cases} \text{reg LIMPt LPIBRt} \\ \text{predict LIMPF,xb} \\ \text{gen LIMPF2=LIMPF}^2 \\ \text{gen LIMPF3=LIMPF}^3 \\ \text{reg LIMPt LPIBRt LIMPF2 LIMPF3} \end{cases}$

Source	SS	df	MS	Number of obs = 34		
Model	1.00781136	3	.335937119	F(3, 30)	=	107.12
Residual	.094081898	30	.003136063	Prob > F	=	0.0000
Total	1.10189325	33	.033390705	R-squared	=	0.9146
				Adj R-squared	=	0.9061
				Root MSE	=	.056
LIMPt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
LPIBRt	-797.2937	223.8832	-3.56	0.001	-1254.524	-340.0632
LIMPF2	179.7882	50.47454	3.56	0.001	76.70547	282.871
LIMPF3	-10.12975	2.846117	-3.56	0.001	-15.94229	-4.317199
_cons	1405.601	393.7628	3.57	0.001	601.4297	2209.771

Commentaire : le modèle est mal spécifié, car tous les paramètres du « modèle 1 » estimés sont statistiquement significatifs. A cela s'ajoute que la statistique de Ramsey-Reset affiche une probabilité $< 5\%$ (on accepte H_1). Autant dire que le PIB réel/PIBR n'est pas la seule variable explicative des importations/IMP, il y a lieu d'intégrer d'autres variables de contrôle.

► **Stabilité des paramètres :**

(i) **Test de CHOW :**

Les hypothèses du test :

H_0 : Le modèle est stable $\rightarrow (prob > 5\%)$

H_1 : Le modèle n'est pas stable $\rightarrow (prob \leq 5\%)$



Pour étudier la stabilité de notre modèle, considérons/définissons deux dates/années de rupture (caractéristiques d'événements particuliers), à savoir : les années 1974 et 1978.

Pour la date/point de rupture « 1974 »

Deux sous-périodes sont à définir :

- o 1^{ère} sous-période : de 1962 à 1973
- o 2^{ème} sous-période : de 1974 à 1995

Sur Eviews, dans l'output, suivre : View/Stability Tests/Chow Breakpoint Test... → 1974 → ok :

Chow Breakpoint Test: 1974			
F-statistic	10.29589	Probability	0.000394
Log likelihood ratio	17.76813	Probability	0.000139

Conclusion : pour les deux sous-périodes considérées, notre modèle estimé n'est pas stable, car : $prob < 5\%$ (les deux probabilités).

Pour la date/point de rupture « 1978 »

Définissons également deux sous-périodes :

- o 1^{ère} sous-période : de 1962 à 1977
- o 2^{ème} sous-période : de 1978 à 1995

Sur Eviews, dans l'output, suivre : View/Stability Tests/Chow Breakpoint Test... → 1978 → ok :

Chow Breakpoint Test: 1978			
F-statistic	1.949199	Probability	0.160004
Log likelihood ratio	4.153795	Probability	0.125318

Conclusion : pour les deux sous-périodes considérées, notre modèle estimé est plutôt stable, car : $prob > 5\%$ (les deux probabilités).

Inconvénient du test de Chow : les résultats par rapport aux deux dates de rupture ne concordent pas (les tests graphiques/CUSUM de Brown, Durbin et Evans tiennent lieu).

Test de Chow sur Stata

Dans Stata, pour calculer la statistique de Chow, procéder comme suit :

- estimer le modèle général ;
- récupérer la somme des carrés des résidus et le nombre d'observations, respectivement « SCR et n » ;
- estimer le modèle 1 pour la 1^{ère} sous-période (1962 – 1977) ;
- récupérer la SCR1 (scalar SCR1=e(rss)) ;
- estimer le modèle 2 pour la 2^{ème} sous-période (1978 – 1995) ;



- récupérer la SCR2 (scalar SCR2=e(rss)) ;
- calcul de la statistique de Chow : $\text{scalar stat} = ((\text{scr} - (\text{scr1} + \text{scr2})) / (\text{scr1} + \text{scr2})) * ((n - 2 * 2) / 2)$;
- display stat : obtenir la valeur calculée ;
- display F(2, n-2*2, stat) : c'est la statistique du test correspondant au « p-value » (1-probabilité du test) de l'hypothèse nulle (H0 : stabilité et H1 : instabilité).

Résultats du test de CHOW :

Le test de CHOW n'est pas programmé dans Stata, mais il peut se calculer comme suit $\left[\frac{\text{SCR} - (\text{SCR1} + \text{SCR2})}{\text{SCR1} + \text{SCR2}} \times \frac{n - 2k}{k} \rightarrow F(k, n - 2k) \right]$; Avec : n (nombre d'observations), k (nombre de paramètres), SCR/SCR1/SCR2 (les sommes des carrés des résidus de deux sous échantillons et de l'échantillon total) :

► $\left\{ \begin{array}{l} \text{reg LIMPt LPIBRt} \\ \text{scalar scr} = \text{e(rss)} \\ \text{scalar n} = \text{e(N)} \end{array} \right\} : \begin{array}{l} \text{Estimation sur l'ensemble de l'échantillon/période et} \\ \text{récupération de « SCR et n »}. \end{array}$

Commande: reg LIMPt LPIBRt

Source	SS	df	MS	Number of obs = 34		
Model	.967660088	1	.967660088	F(1, 32)	=	230.68
Residual	.134233167	32	.004194786	Prob > F	=	0.0000
Total	1.10189325	33	.033390705	R-squared	=	0.8782
				Adj R-squared	=	0.8744
				Root MSE	=	.06477
LIMPt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
LPIBRt	.7506297	.0494219	15.19	0.000	.6499607	.8512988
_cons	.6487534	.3469917	1.87	0.071	-.0580455	1.355552

scalar scr=e(rss)
scalar n=e(N)

► $\left\{ \begin{array}{l} \text{reg LIMPt LPIBRt if annee} \leq 1977 \\ \text{scalar scr1} = \text{e(rss)} \end{array} \right\} : \begin{array}{l} \text{Estimation sur le 1}^{\text{er}} \text{ sous-échantillon} \\ \text{(1962-1977) et récupération de SCR1.} \end{array}$

Commande: reg LIMPt LPIBRt if annee<=1977

Source	SS	df	MS	Number of obs = 16		
Model	.076458508	1	.076458508	F(1, 14)	=	10.57
Residual	.101307768	14	.007236269	Prob > F	=	0.0058
Total	.177766276	15	.011851085	R-squared	=	0.4301
				Adj R-squared	=	0.3894
				Root MSE	=	.08507
LIMPt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
LPIBRt	.6287474	.1934286	3.25	0.006	.2138844	1.04361
_cons	1.467453	1.318822	1.11	0.285	-1.36114	4.296045

scalar scr1=e(rss)



► $\left\{ \begin{array}{l} \text{reg LIMPt LPIBRt if annee} > 1977 \\ \text{scalar scr2} = \text{e(rss)} \end{array} \right. : \left\{ \begin{array}{l} \text{Estimation sur le 2}^{\text{ème}} \text{ sous-échantillon} \\ (1978-1995) \text{ et récupération de SCR2.} \end{array} \right.$

Commande: reg LIMPt LPIBRt if annee>1977

Source	SS	df	MS	Number of obs = 18		
Model	.109041999	1	.109041999	F(1, 16)	=	99.76
Residual	.017488168	16	.001093011	Prob > F	=	0.0000
Total	.126530167	17	.007442951	R-squared	=	0.8618
				Adj R-squared	=	0.8531
				Root MSE	=	.03306
LIMPt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
LPIBRt	.5901453	.0590846	9.99	0.000	.4648916	.7153991
_cons	1.81434	.4252052	4.27	0.001	.9129456	2.715735

scalar scr2=e(rss)

► $\left\{ \begin{array}{l} \text{scalar stat} = ((\text{scr} - (\text{scr1} + \text{scr2})) / (\text{scr1} + \text{scr2})) * ((n - 2 * 2) / 2) \\ \text{display stat} \\ \text{display F}(2, n - 2 * 2, \text{stat}) \end{array} \right.$

NB :

- o 1^{ère} expression (scalar...) : calculer la statistique de Fisher $F(2, n-2*2)$;
- o 2^{ème} expression (display stat) : afficher la valeur F-calculé : **stat = 1.9492119. Le F-théorique/F(0.05 ; 2 ; 12) = 3.89;**
- o 3^{ème} expression (display F(2,n-2*2,stat)) : calculer la probabilité associée à la statistique calculée de Fisher (**NB** : par définition, la p-value est le complément à 1 de la probabilité du test. C.à.d. « p-value = 1-prob du test »).
prob F (p-value nulle) = 0.83999769, et p-value = 1-0.83999769=0.16.

L'on constate ainsi que la « p-value » associée au F-stat est > à 5% ($F_c < F_t$, soit **1.9492119 < 3.89**). D'où, on accepte l'hypothèse nulle selon la quelle les paramètres estimés sont stables sur les deux sous-périodes/les deux régressions ne sont pas différentes.

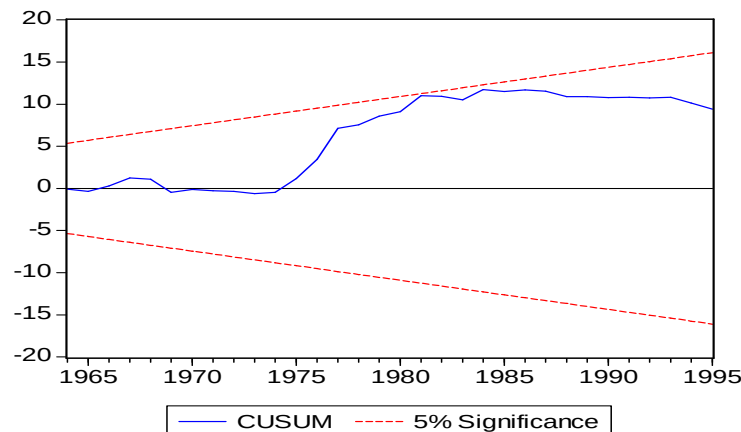
(ii) **Test de CUSUM (Brown, Durbin et Evans) :**

L'hypothèse de stabilité du modèle est validée si la courbe Cusum ne sort pas du corridor (intervalle de confiance). Au cas contraire, l'on parlera de l'instabilité structurelle (test Cusum) ou l'instabilité ponctuelle (test Cusum carré). Pour ces deux tests, seul Eviews a été utilisé.

Résultat du test CUSUM

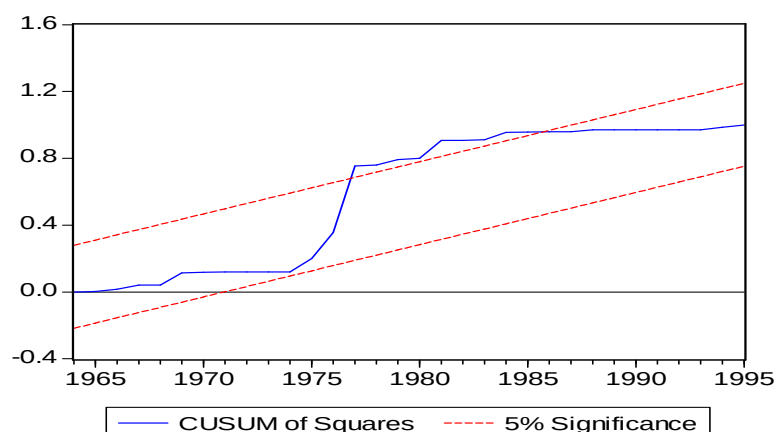
Sur Eviews, dans l'output, suivre : View/Stability Tests/Recursive Estimates (OLS only)... → Cusum Test → ok : La courbe ne coupe pas le corridor en pointillé ; d'où, le modèle est structurellement stable.





📊 Résultat du test Cusum Carré

Sur Eviews, dans l'output, suivre : View/Stability Tests/Recursive Estimates (OLS only)... → Cusum of Squares Test → ok : La courbe sort du corridor en pointillé ; d'où, le modèle souffre d'une instabilité ponctuelle entre 1977 et 1986, éventuellement expliquée par le 2^{ème} choc pétrolier, sans se passer des politiques d'ajustement structurel au Sénégal.



(iii) Stabilisation du modèle

Le recours aux variables dummy (indicatrices, muettes, etc.) est à compter parmi les techniques de stabilisation d'un modèle. Dans notre cas, nous créons la variable dummy/dichotomique « *instab* » avec les propriétés suivantes :

$$instab = \begin{cases} 1, & \text{si année} \in [1977, 1986] \\ 0, & \text{ailleurs} \end{cases}$$

____ **Sur Eviews, taper :**

```
{ Genr instab=0*(annee<1977)+0*(annee>1986)+1*(annee>=1977 and annee<=1986)
  Ls limp c lpibr instab
```

____ **Sur Stata, taper :**

```
{ gen instab=1
  replace instab=0 if annee<1977
  replace instab=0 if annee>1986
  reg LIMPt LPIBRt instab
```



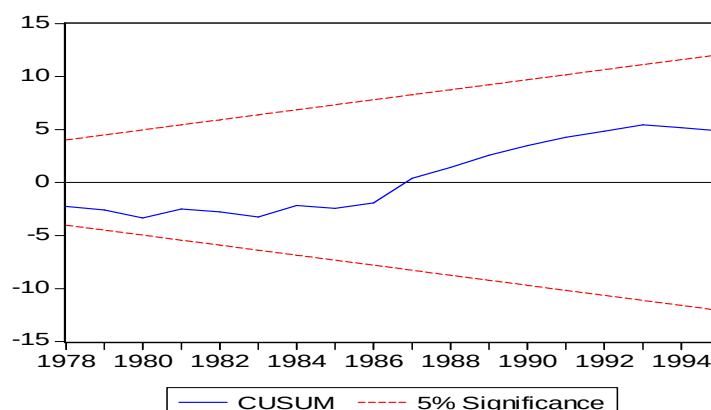
Dependent Variable: LIMP				
Method: Least Squares				
Date: 12/01/13 Time: 01:48				
Sample (adjusted): 1962 1995				
Included observations: 34 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.768366	0.319580	2.404298	0.0224
LPIBR	0.731009	0.045657	16.01081	0.0000
INSTAB	0.061461	0.022521	2.729113	0.0104
R-squared	0.901778	Mean dependent var	5.916234	
Adjusted R-squared	0.895441	S.D. dependent var	0.182731	
S.E. of regression	0.059087	Akaike info criterion	-2.735510	
Sum squared resid	0.108230	Schwarz criterion	-2.600832	
Log likelihood	49.50368	F-statistic	142.3061	
Durbin-Watson stat	1.174653	Prob(F-statistic)	0.000000	

Commande: reg LIMPt LPIBRt instab

Source	SS	df	MS	Number of obs = 34		
Model	.993663437	2	.496831719	F(2, 31)	=	142.31
Residual	.108229817	31	.003491284	Prob > F	=	0.0000
Total	1.10189325	33	.033390705	R-squared	=	0.9018
				Adj R-squared	=	0.8954
				Root MSE	=	.05909
LIMPt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
LPIBRt	.7310089	.0456572	16.01	0.000	.6378904	.8241273
instab	.0614611	.0225205	2.73	0.010	.0155302	.1073921
_cons	.7683645	.3195801	2.40	0.022	.1165766	1.420152

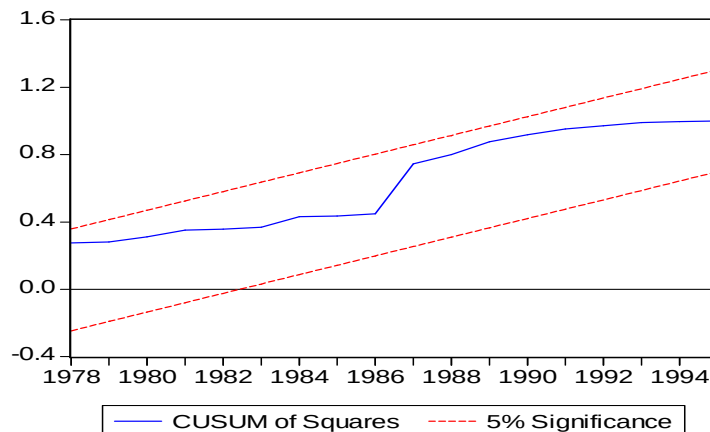
Commentaire : le 2^{ème} choc pétrolier et les politiques d'ajustement structurel ont été favorables aux importations du Sénégal en ce qu'ils l'ont augmentés de l'ordre de 0.06 point.

Test CUSUM : Sur Eviews, dans l'output, suivre: View/Stability Tests/Recursive Estimates (OLS only)... → Cusum Test → ok : La courbe ne coupe pas le corridor en pointillé; d'où, le modèle stabilisé est toujours structurellement stable.



Test Cusum Carré : Sur Eviews, dans l'output, suivre: View/Stability Tests/Recursive Estimates (OLS only)... → Cusum of Squares Test → ok : La courbe évolue maintenant dans les limites du corridor en pointillé; d'où, le modèle stabilisé est ponctuellement stable.

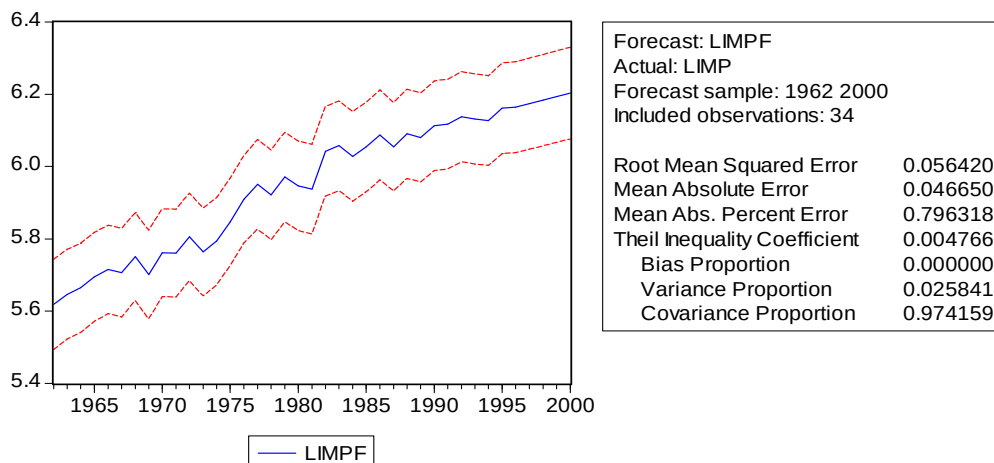




d) Prédiction des importations pour les années 1996, 1997, 1998, 1999 et 2000

Notre modèle étant stable, faisons la prédiction pour les cinq années comme suit :

Sur Eviews, dans l'out put du modèle stabilisé (avec « instab »), cliquer sur **Forecast** → taper : `limpf` → ok :



Ensuite, faire : `genr impf=exp(limpf)`

Avec : « `limpf` » : la variable « `limp` » prévue, et « `impf` » : la variable « `imp` » prévue (l'exponentiel de la variable logarithmique « `limp` »). Les valeurs prévues sont consignées dans le tableau ci-dessous :

Années	LIMPF	IMPF
1996	6.16430763811427	475.471830373706
1997	6.1742085839334	480.202833258587
1998	6.18393290889367	484.895260020134
1999	6.19357330369756	489.592446753993
2000	6.20308821596185	494.273108660778

Commentaire : la prédiction est bonne au regard – à titre illustratif – de l'indice d'inégalité de Theil qui avoisine zéro (proche de 0) et du MAPE = 0.79% (Mean Absolute Percent Error).



Sur Stata, les commandes sont :

```

gen instab=1
replace instab=0 if annee<1977
replace instab=0 if annee>1986
reg LIMPt LPIBRt instab
predict LIMPF, xb
gen IMPFt=exp(LIMPF)
list LIMPF IMPFt if annee>1995

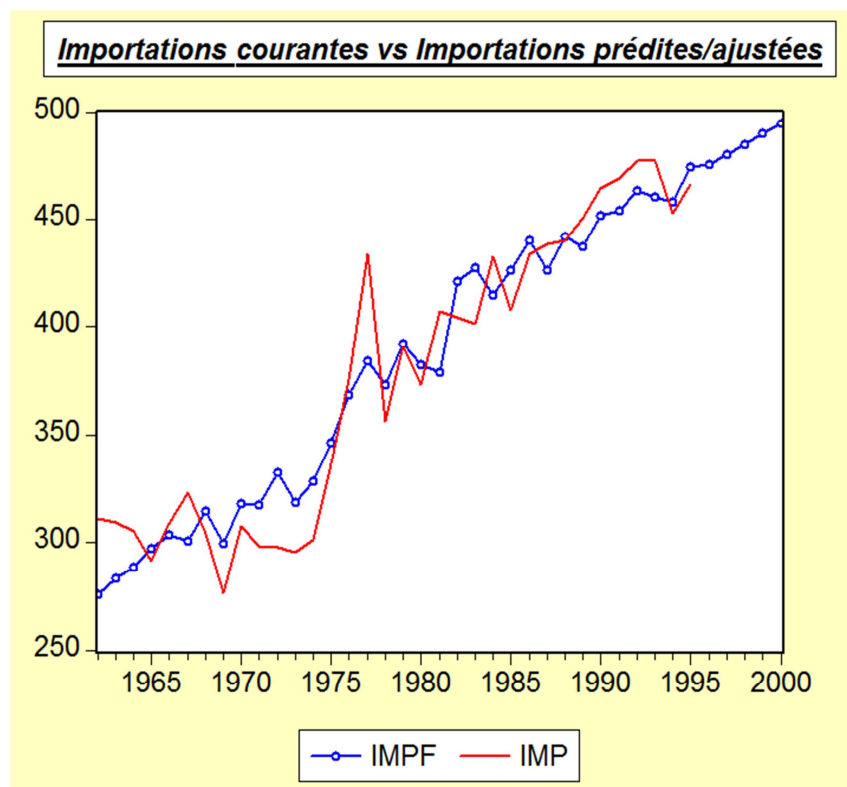
```

Les valeurs prévues sous Stata se présentent comme suit :

	LIMPF	IMPFt
35.	6.164308	475.4718
36.	6.174209	480.2029
37.	6.183933	484.8952
38.	6.193573	489.5925
39.	6.203088	494.2731

Graphiquement, les valeurs courantes et ajustées/prédites des importations (IMP) se présentent comme suit :

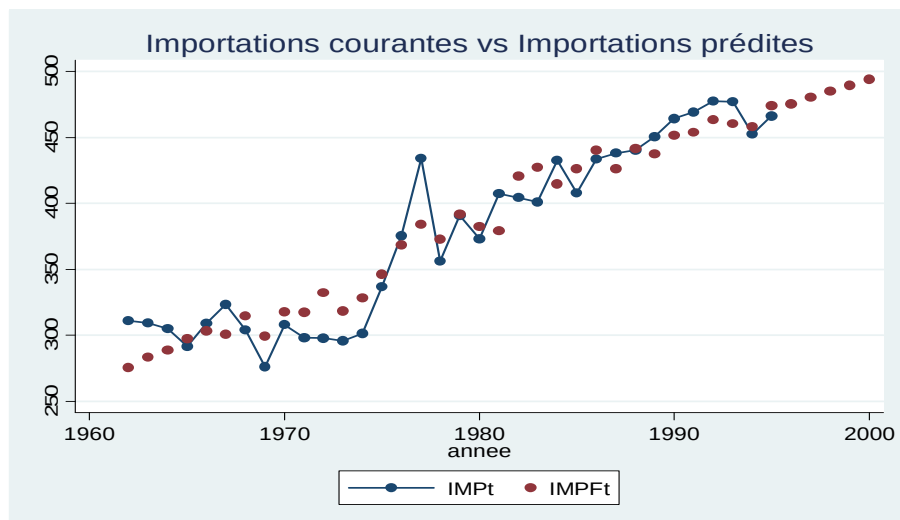
o Sur Eviews, taper : plot impf imp



o Sur Stata :

Commande : scatter IMPt IMPF annee, connect(.l)
title(Importations courantes vs Importations prédites)





CHAP II MODELES A DECALAGES TEMPORELS

Cas pratique 2.1 : Investissement fonction du PIB et du taux d'intérêt

► **Modèle :**

$$\log(INV_t) = b_0 + b_1 \log(INV_{t-1}) + b_2 \log(PIBR_t) + b_3 TI_t + e_t \dots \dots [2.1]$$

_____ Avec : INV_t = investissement courant ; INV_{t-1} = investissement passé (1 an passé) ; $PIBR_t$ = PIB réel courant ; et TI_t = taux d'intérêt réel courant.

► **Travail demandé (recours à Stata et Eviews) :**

- o Estimer les paramètres par les MCO et interpréter le R^2 ;
- o Effectuer les tests usuels [significativité individuelle (Student) et globale des paramètres (Fisher), hétéroscédasticité (White), bonne spécification (test reset de Ramsey) et stabilité des coefficients (test CUSUM)], celui d'autocorrélation des erreurs de Breusch-Godfrey et le test h de Durbin (test d'autocorrélation des erreurs adapté pour les modèles dynamiques) ;
- o Prévoir l'investissement pour les années 2002, 2003 et 2004.

► **Résolution :**

a) Estimation des paramètres par les MCO et interprétation du R^2 :

_____ Commandes Stata :

```
{ tsset temps
  gen LINVt=log(INVt)
  gen LPIBRt=log(PIBRt)
```

_____ Commandes Eviews :

```
{ create a 1972 2004
  data INVt PIBRt Tit
  ls log(INVt) c log(INVt(-1)) log(PIBRt) TIt
```



Dependent Variable: LOG(INVT)				
Method: Least Squares				
Date: 12/07/13 Time: 15:18				
Sample (adjusted): 1973 2001				
Included observations: 29 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.265328	0.398106	-3.178371	0.0039
LOG(INVT(-1))	0.428072	0.083538	5.124291	0.0000
LOG(PIBRT)	0.646582	0.102471	6.309894	0.0000
TIT	-0.047420	0.009921	-4.779499	0.0001
R-squared	0.965956	Mean dependent var	6.183779	
Adjusted R-squared	0.961871	S.D. dependent var	0.501060	
S.E. of regression	0.097841	Akaike info criterion	-1.683511	
Sum squared resid	0.239320	Schwarz criterion	-1.494918	
Log likelihood	28.41091	F-statistic	236.4473	
Durbin-Watson stat	2.403637	Prob(F-statistic)	0.000000	

Commande: `reg LINVt L.LINVt LPIBRT TIT`

Source	SS	df	MS	Number of obs = 29		
Model	6.79038454	3	2.26346151	F(3, 25) =	236.45	
Residual	.239319785	25	.009572791	Prob > F =	0.0000	
Total	7.02970433	28	.251060869	R-squared =	0.9660	
				Adj R-squared =	0.9619	
				Root MSE =	.09784	
LINVt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
LINVt						
L1.	.4280725	.0835378	5.12	0.000	.2560232	.6001219
LPIBRT	.6465819	.1024711	6.31	0.000	.4355386	.8576251
TIT	-.0474197	.0099215	-4.78	0.000	-.0678534	-.026986
_cons	-1.265328	.3981059	-3.18	0.004	-2.085243	-.4454139

Interprétation du R^2 : le modèle autorégressif d'ordre 1 ---- tel que spécifié (variables explicatives/investissement, relation linéaire, etc.) ---- explique 96,59% de variations/évolutions des investissements ($R^2 = 0.9659$). Autant dire que les investissements antérieurs (il y a une année passée), le PIB réel et le taux d'intérêt courants contribuent à expliquer 96,59% de variations intervenues dans les investissements courants.

b) Tests usuels (classiques)

- ✚ **Significativité individuelle**: tous les paramètres estimés « b_0, b_1, b_2 et b_3 » sont statistiquement significatifs au seuil de 1% (au regard des probabilités statistiques associées aux t de student calculés);
- ✚ **Significativité conjointe**: Pris ensemble, ces paramètres estimés sont statistiquement significatifs au regard de la statistique calculée de Fisher (prob F-stat < 5%);
- ✚ **Test d'homoscédasticité de White**: Les erreurs sont homoscédastiques (pour le test sans termes croisés/no cross terms: résultat à gauche), car: prob-F > 5% et tous les paramètres estimés sont statistiquement non significatifs. Par contre, pour les cross terms (à droite), les erreurs sont hétéroscédastiques (résultat adopté: envisager une correction au sens de White).

Les hypothèses du test :

$$\begin{cases} H_0 : \text{Homoscédasticité (prob-F > 5\%)} \\ H_1 : \text{Absence d'homoscédasticité (prob-F < 5\%)} \end{cases}$$



Commandes Stata :

```

reg LINVt L.LINVt LPIBRT TIt
predict e, resid
gen e2=e^2
gen LINVtd=LINVt[_n]
gen LINVtd2=LINVt^2
gen LPIBRT2=LPIBRT^2
gen TIt2=TIt^2
reg e2 LINVtd LINVtd2 LPIBRT LPIBRT2 TIt TIt2

```

Source	SS	df	MS	Number of obs = 29		
Model	.001154376	6	.000192396	F(6, 22) =	1.24	
Residual	.003425937	22	.000155724	Prob > F =	0.3265	
Total	.004580313	28	.000163583	R-squared =	0.2520	
				Adj R-squared =	0.0480	
				Root MSE =	.01248	
e2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
LINVtd	-.0273308	.3197778	-0.09	0.933	-.6905095	.6358478
LINVtd2	.0012648	.0255829	0.05	0.961	-.0517909	.0543205
LPIBRT	-.2999637	.2751096	-1.09	0.287	-.8705061	.2705787
LPIBRT2	.0195299	.0169924	1.15	0.263	-.0157101	.05477
TIt	-.0130348	.0159401	-0.82	0.422	-.0460926	.020023
TIt2	.000812	.0008492	0.96	0.349	-.0009492	.0025732
_cons	1.321499	1.128384	1.17	0.254	-1.018626	3.661624

Sur Eviews, dans l'output, suivre: View/Residual Tests/White Heteroskedasticity (no cross terms ou cross terms) → ok :

White Heteroskedasticity Test:				
F-statistic	1.640577	Probability	0.183271	
Obs*R-squared	8.964491	Probability	0.175585	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 12/07/13 Time: 17:23				
Sample: 1973 2001				
Included observations: 29				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.188103	0.983680	1.207814	0.2399
LOG(INVT(-1))	0.329588	0.258699	1.274024	0.2159
(LOG(INVT(-1)))^2	-0.027125	0.020812	-1.303364	0.2059
LOG(PIBRT)	-0.525989	0.322083	-1.633086	0.1167
(LOG(PIBRT))^2	0.032786	0.019712	1.663245	0.1104
TIT	-0.017641	0.013511	-1.305699	0.2051
TIT^2	0.001037	0.000733	1.413697	0.1714
R-squared	0.309120	Mean dependent var	0.008252	
Adjusted R-squared	0.120699	S.D. dependent var	0.012790	
S.E. of regression	0.011993	Akaike info criterion	-5.802438	
Sum squared resid	0.003164	Schwarz criterion	-5.472401	
Log likelihood	91.13534	F-statistic	1.640577	
Durbin-Watson stat	2.339800	Prob(F-statistic)	0.183271	

White Heteroskedasticity Test:				
F-statistic	4.627231	Probability	0.002410	
Obs*R-squared	19.91435	Probability	0.018448	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 12/07/13 Time: 17:24				
Sample: 1973 2001				
Included observations: 29				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-4.999222	1.684518	-2.967746	0.0079
LOG(INVT(-1))	0.403888	0.316475	1.276209	0.2173
(LOG(INVT(-1)))^2	-0.037876	0.024616	-1.538686	0.1404
(LOG(INVT(-1)))*(LOG(PIB...	0.015877	0.059367	0.267445	0.7920
(LOG(INVT(-1))*TIT	-0.007280	0.006777	-1.074198	0.2962
LOG(PIBRT)	1.185898	0.570859	2.077393	0.0516
(LOG(PIBRT))^2	-0.095722	0.054648	-1.751588	0.0960
(LOG(PIBRT))*TIT	0.029087	0.009176	3.169875	0.0050
TIT	-0.221819	0.044740	-4.957918	0.0001
TIT^2	0.001610	0.000724	2.225424	0.0384
R-squared	0.686702	Mean dependent var	0.008252	
Adjusted R-squared	0.538297	S.D. dependent var	0.012790	
S.E. of regression	0.008691	Akaike info criterion	-6.386351	
Sum squared resid	0.001435	Schwarz criterion	-5.914870	
Log likelihood	102.6021	F-statistic	4.627231	
Durbin-Watson stat	2.622795	Prob(F-statistic)	0.002410	



Tests d'autocorrélation des erreurs :

• Test h de Durbin :

Dans le cas de modèles dynamiques, pour tester l'autocorrélation des erreurs, l'on recourt à la statistique h de Durbin⁽¹⁾ exprimée comme suit :

$$h = \left(1 - \frac{d^*}{2}\right) \sqrt{\frac{T}{1 - T[\text{var}(\hat{b}_1)]}} = \hat{\rho} \sqrt{\frac{T}{1 - T[\text{var}(\hat{b}_1)]}}$$

$$= \left(1 - \frac{2,4036}{2}\right) \sqrt{\frac{29}{1 - 29[(0,083538)^2]}} \rightarrow h = -1,216$$

Avec :

- T : Taille de l'échantillon/nombre d'observations ;
- $\text{var}(\hat{b}_1)$: variance du coefficient, associé à INVt-1 , estimé ;
- d^* : statistique de Durbin-Watson calculée ($h \sim \text{loi normale}$ pour un échantillon grande de taille). NB : sur stata, faire : **dwstat**

Durbin-Watson d-statistic(4, 29) = 2.403636
--

- $\hat{\rho} = 1 - \frac{d^*}{2}$.

Les hypothèses du test sont :

H_0 : absence d'autocorrélation ($h < z$ ou $h > 1,96$; $\text{prob} > 5\%$)

H_1 : Présence d'autocorrélation ($h \geq z$ ou $h \leq -1,96$; $\text{prob} < 5\%$)

L'on notera que sous H_0 , « h » est normalement distribuée (« z » est la valeur lue sur la table de la loi normale : $\alpha = 5\%$ et $z = \pm 1,96$).

Conclusion : Suivant la table de la loi normale, $|h| < z$ (d'où, accepter H_0 : absence d'autocorrélation de résidus au seuil de 5%).

- Test de Breush-Godfrey : les erreurs ne sont pas auto-corrélées à l'ordre 2 (ni à l'ordre 1), car : $\text{prob} > 5\%$.

Les hypothèses du test :

H_0 : Absence d'autocorrélation ($\text{prob} > 5\%$)

H_1 : Présence d'autocorrélation ($\text{prob} < 5\%$)

Sur stata, obtenir :

Commande: `bgodfrey, lags(2) small`

Breusch-Godfrey LM test for autocorrelation

lags(p)	F	df	Prob > F
2	1.352	(2, 23)	0.2786
H0: no serial correlation			

¹ Le test de Durbin-Watson traditionnel sous-estime le risque d'autocorrélation dans ce type de modèle : il n'est donc pas efficace.



____ Sur Eviews, dans l'output, suivre : View/Residual Tests/Serial correlation LM Test... → lag to include : 1 → ok :

Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	1.182182	Probability	0.324560	
Obs*R-squared	2.703263	Probability	0.258818	
Test Equation:				
Dependent Variable: RESID				
Method: Least Squares				
Date: 12/07/13 Time: 17:17				
Presample missing value lagged residuals set to zero.				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.026593	0.434483	-0.061205	0.9517
LOG(INVT(-1))	0.000870	0.099039	0.008788	0.9931
LOG(PIBRT)	0.003572	0.120594	0.029622	0.9766
TIT	-0.000866	0.009978	-0.086745	0.9316
RESID(-1)	-0.170232	0.236687	-0.719229	0.4792
RESID(-2)	0.222638	0.226949	0.981004	0.3368
R-squared	0.093216	Mean dependent var	-1.41E-15	
Adjusted R-squared	-0.103911	S.D. dependent var	0.092451	
S.E. of regression	0.097135	Akaike info criterion	-1.643431	
Sum squared resid	0.217011	Schwarz criterion	-1.360542	
Log likelihood	29.82975	F-statistic	0.472873	
Durbin-Watson stat	1.904235	Prob(F-statistic)	0.792568	

🌈 Test de bonne spécification Reset de Ramsey :

____ Les hypothèses du test :

H_0 : Le modèle est bien spécifié → ($prob > 5\%$)

H_1 : Le modèle n'est pas bien spécifié → ($prob < 5\%$)

Commandes Stata: $\begin{cases} \text{reg LINvt L.LINvt LPIBrt Tit} \\ \text{ovtest} \end{cases}$

Ramsey RESET test using powers of the fitted values of LINvt
 H_0 : model has no omitted variables
 F(3, 22) = 2.33
 Prob > F = 0.1022

____ Sur Eviews, dans l'output, suivre : View/Stability Test/Ramsey reset Test/Number of fitted term → 2 → ok :

Ramsey RESET Test:				
F-statistic	2.158738	Probability	0.138294	
Log likelihood ratio	4.988941	Probability	0.082540	
Test Equation:				
Dependent Variable: LOG(INVT)				
Method: Least Squares				
Date: 12/07/13 Time: 18:46				
Sample: 1973 2001				
Included observations: 29				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	83.37132	81.16797	1.027145	0.3150
LOG(INVT(-1))	-10.30441	10.63231	-0.969160	0.3425
LOG(PIBRT)	-15.48694	16.15672	-0.958545	0.3478
TIT	1.148140	1.180599	0.972507	0.3409
FITTED^2	3.896888	4.172005	0.934057	0.3600
FITTED^3	-0.201369	0.232232	-0.867103	0.3948
R-squared	0.971337	Mean dependent var	6.183779	
Adjusted R-squared	0.965105	S.D. dependent var	0.501060	
S.E. of regression	0.093599	Akaike info criterion	-1.717612	
Sum squared resid	0.201496	Schwarz criterion	-1.434723	
Log likelihood	30.90538	F-statistic	155.8829	
Durbin-Watson stat	2.879809	Prob(F-statistic)	0.000000	

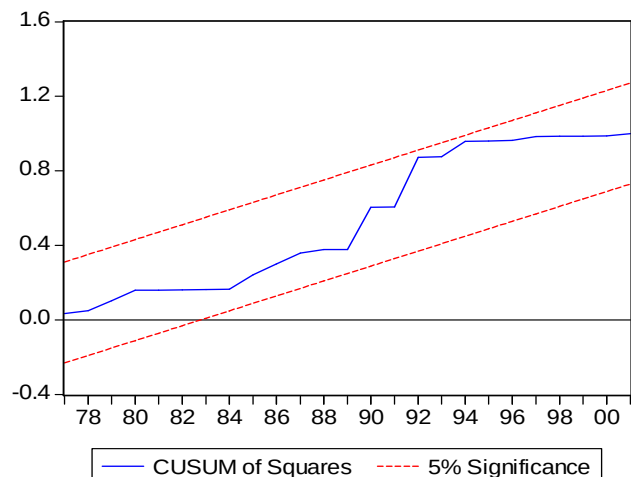
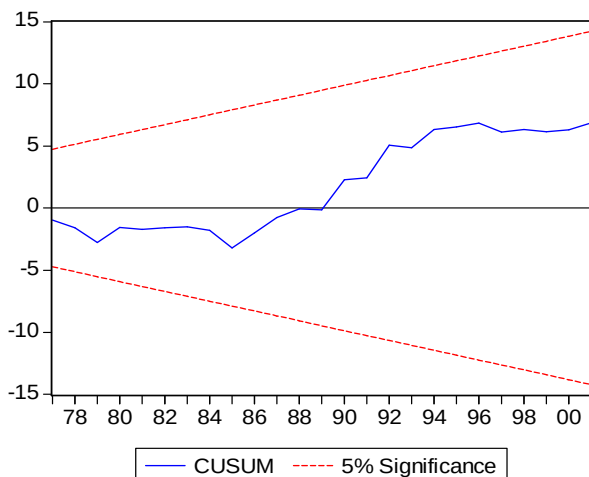


Commentaire : le modèle est bien spécifié, car tous les paramètres estimés du modèle contraint (Cfr output Eviews) sont statistiquement non significatifs. A cela s'ajoute que la statistique de Ramsey-Reset affiche une probabilité $> 5\%$ (on accepte H_0).

Test de stabilité de paramètres

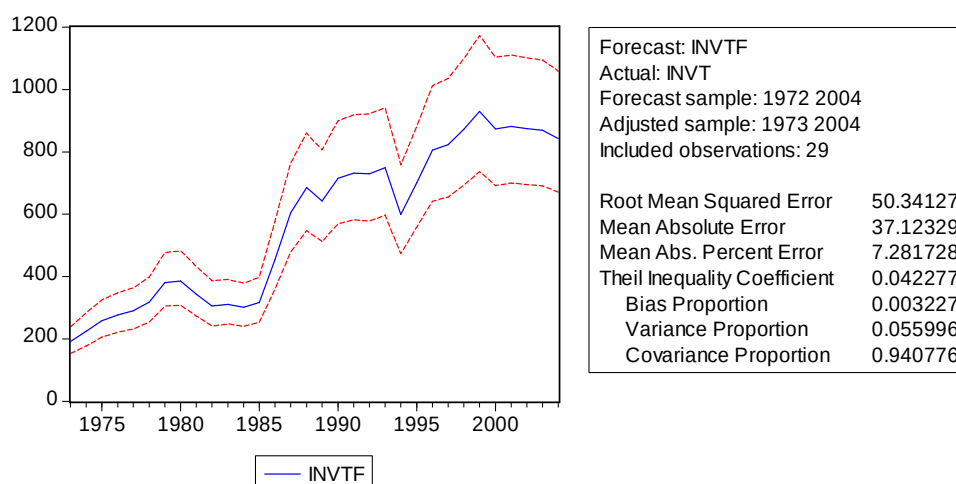
Sur Eviews, suivre les chemins :

- View/stability Tests/Recursive Estimates/Cusum Test → ok : le modèle est structurellement stable (Cfr graphique à gauche) ;
- View/stability Tests/Recursive Estimates/Cusum of Squares Test → ok : le modèle est ponctuellement stable (Cfr graphique à droite).



c) **Prévision de l'investissement pour les années 2002, 2003 et 2004**

Sur Eviews, dans l'output, cliquer sur **Forecast** → ok :

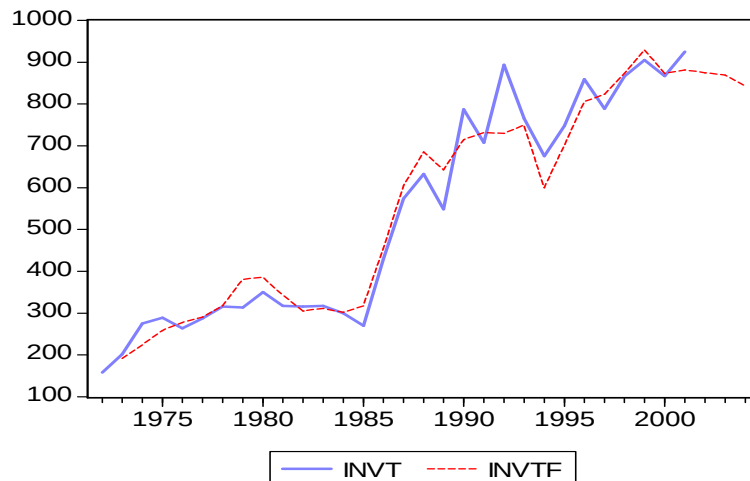


- ❖ **Constat :** MAPE = 7,28% et indice de Theil = 0.04 (notre modèle prédit avec efficacité les valeurs prises par l'investissement) ;
- ❖ Pour voir les valeurs prédites et prévues (avec INVtf : variable INVT estimée/ajustée ou prédite), sur Eviews, taper : **show INVtf** :



Années	Valeurs prévues
2002	874.991402620589
2003	869.355699857323
2004	842.397997064215

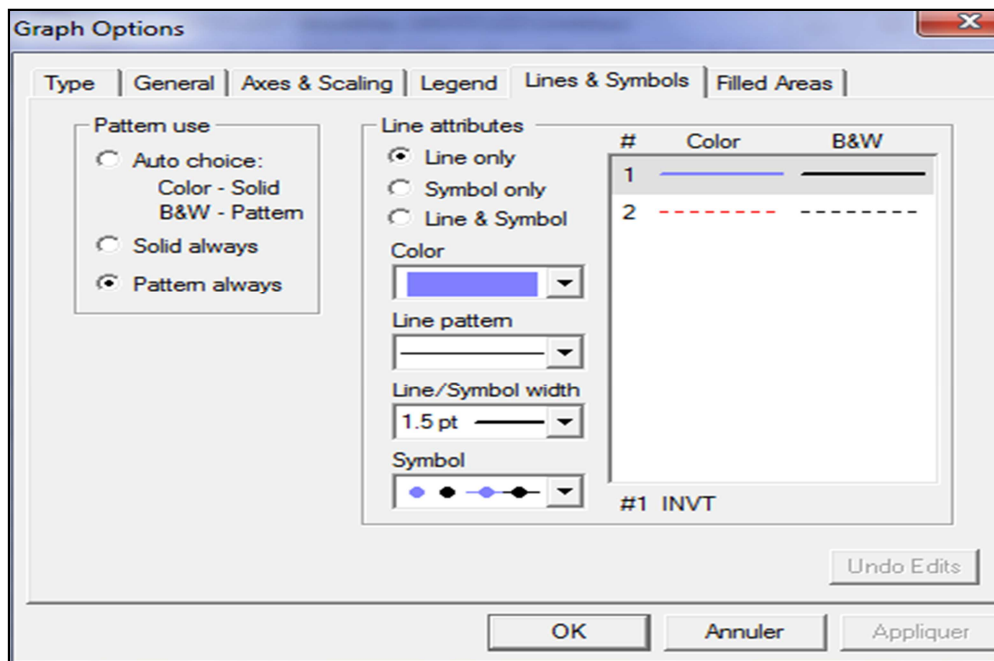
- ❖ Sur Eviews, pour visualiser le graphique des valeurs réelles et prévues, dans l'output, taper : **plot INVT INVTf** :



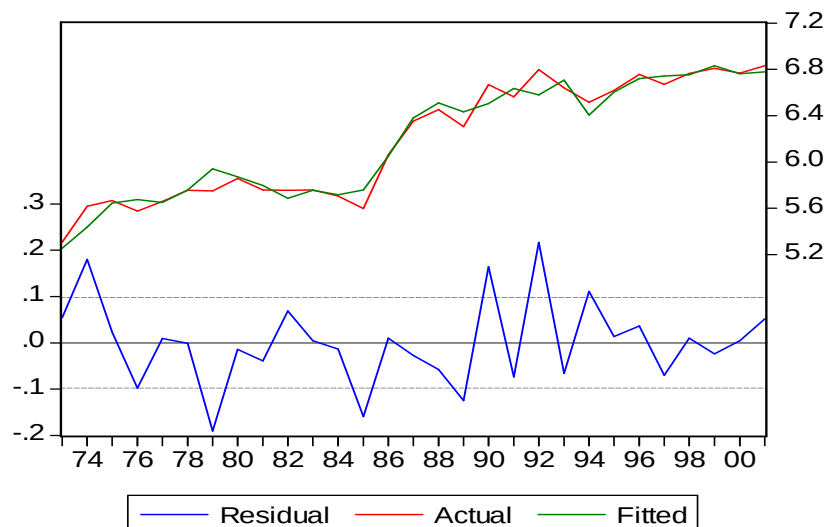
- ❖ Sur Eviews, dans l'output du graphique, cliquer sur **Options** pour modifier le graphique ci-dessus comme suit :



_____ Pour ce faire, Cfr les options de la boîte de dialogue « Graph Options » suivante (Type, General, Axes & Scaling, Legend, Line & Symbols, Filled Areas) :



- ❖ Sur Eviews, pour visualiser le graphique des valeurs réelles et prévues (simulation avec les résidus), dans l'output, cliquer sur **Resids** :



Avec :

- o $\text{Log}(\text{INVt}) = \text{Actual}$;
- o $\text{Log}(\text{INVt}) \text{ estimé} = \text{Fitted}$; et
- o $\text{Residual} = [\text{log}(\text{INVt}) - \text{log}(\text{INVt}) \text{ estimé}]$



Cas pratique 2.2 : Dépenses d'Investissement fonction des Profits passés

► Modèle :

$$I_t = b_0 + b_1 P_t + b_2 P_{t-1} + b_3 P_{t-2} + \dots + b_h P_{t-h} + e_t \dots \dots [2.2]$$

Avec : I_t = dépenses en investissement courant ; P_t = profits courants ; et P_{t-h} = profits antérieurs (pour h périodes).

- **Hypothèse** : les profits antérieurs (P_{t-h}) contribuent à expliquer les dépenses en investissement (I_t) : Cfr théorie économique.
- **Travail demandé (recours à Stata et Eviews)** : déterminer le nombre de décalages optimal (trimestriel) au bout duquel les effets des variations des profits se font ressentir sur les dépenses en investissement et estimer le modèle à retards échelonnés adéquat (par principe de parcimonie).

► Résolution :

a) Déclaration de données trimestrielles à Stata :

Taper `edit/Saisir` les codes numériques du 1^{er} trimestre 1980 au dernier trimestre 1990 (en commençant par 80, avec « 1 » comme raison. **NB** : 1980-1960 = 20 x 4 trimestres par moi = 80). Après avoir fermé le data editor, taper : **format var3 %tq**. Ensuite, taper : **tsset var3** : le message suivant s'affiche :

time variable : var3, 1980q1 to 1990q4

Les données se présentent en partie comme suit :

	It	Pt	var3
1	2072	1660	1980q1
2	2077	1926	1980q2
3	2078	2181	1980q3
4	2043	1897	1980q4
5	2062	1695	1981q1
6	2067	1705	1981q2
7	1964	1731	1981q3
8	1981	2151	1981q4
9	1914	2556	1982q1
10	1991	3152	1982q2
11	2129	3763	1982q3
12	2309	3903	1982q4
13	2614	3912	1983q1
14	2896	3571	1983q2

b) Détermination du décalage optimal

Pour ce faire, calculer/obtenir les critères de Akaike et Schwarz, pour un retard de tâtonnement de 10 ($44/4 = 11$), en estimant les modèles à retards échelonnés d'ordre 1 jusqu'à 10 (retenir le modèle qui minimise ces critères). Les résultats des critères calculés sont (le décalage optimal retenu est $h = 6$) :



Retard/décalage	Akaike	Schwarz
0	14.88	14.96
1	14.42	14.55
2	13.97	14.14
3	13.48	13.69
4	13.18	13.44
5	12.93	13.23
6	12.78	13.13
7	12.83	13.22
8	12.91	13.35
9	12.98	13.47
10	13.05	13.59

c) *Estimation du modèle à retards échelonnés*

Commande : reg It Pt L.Pt L2.Pt L3.Pt L4.Pt L5.Pt L6.Pt						
Source	SS	df	MS	Number of obs = 38		
Model	6155805.34	7	879400.763	F(7, 30) = 50.63		
Residual	521116.052	30	17370.5351	Prob > F = 0.0000		
Total	6676921.39	37	180457.335	R-squared = 0.9220		
				Adj R-squared = 0.9037		
				Root MSE = 131.8		
It	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
Pt						
--.	-.011389	.0815322	-0.14	0.890	-.1779	.1551221
L1.	.061265	.1249063	0.49	0.627	-.1938278	.3163577
L2.	.2275692	.1196352	1.90	0.067	-.0167584	.4718968
L3.	.1679316	.1129971	1.49	0.148	-.0628393	.3987026
L4.	.1187338	.1274542	0.93	0.359	-.1415624	.3790301
L5.	.0001691	.1369068	0.00	0.999	-.2794319	.2797702
L6.	.2371737	.0840651	2.82	0.008	.0654899	.4088575
_cons	501.5414	154.8486	3.24	0.003	185.2984	817.7845

Commentaire : Les effets positifs de l'accumulation des profits antérieurs se font ressentir sur les dépenses en investissement après 1 an et demi (soit 6 trimestres : seul le coefficient du 6^{ème} retard est statistiquement significatif, y compris la constante).



CHAP III :

COINTEGRATION, MODELE A CORRECTION D'ERREUR/MCE ET VECTEUR A CORRECTION D'ERREUR/VECM

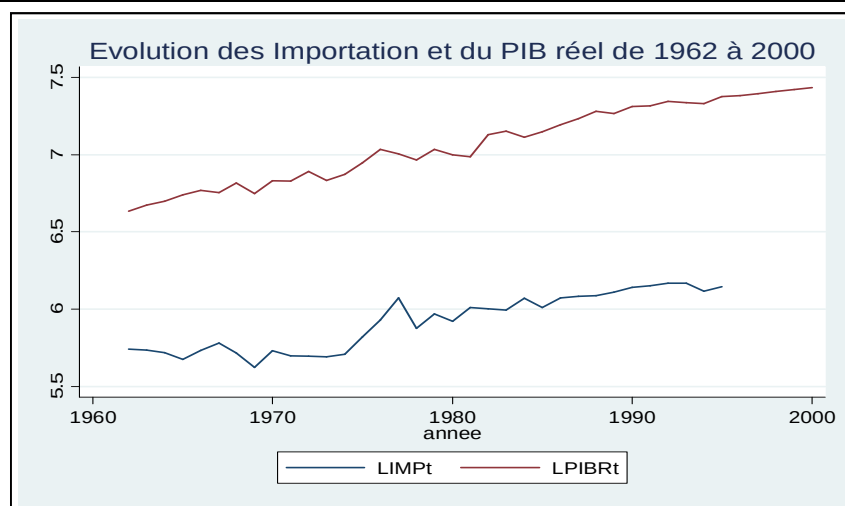
A. ESTIMATION D'UN MCE

Cas pratique 3.1: Test de stationnarité, test de cointégration et estimation d'un MCE : Importation fonction du PIB réel

III.1. Graphique et corrélogrammes

_____ Evolution graphique des variables : sur Stata, la commande est :

```
(line LIMPt LPIBRt annee),title(Evolution des Importation et du PIB réel de 1962 à 2000)
```



Note: les séries « limpt et lpibr » ont une tendance commune (elles semblent cointégrées).

_____ Corrélogrammes des variables :

- Sur Eviews, suivre : Quick/Series Statistics/Correlogram... → taper **limpt** ou **lpibr** → ok → cocher **Level** → ok ou taper **ident limpt** :

Date: 12/09/13 Time: 12:57 Sample: 1962 2000 Included observations: 34					
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.892	0.892	29.497	0.000	
2	0.829	0.165	55.787	0.000	
3	0.738	-0.133	77.285	0.000	
4	0.661	-0.027	95.091	0.000	
5	0.584	-0.014	109.49	0.000	
6	0.454	-0.325	118.50	0.000	
7	0.351	-0.053	124.07	0.000	
8	0.214	-0.177	126.24	0.000	
9	0.143	0.160	127.24	0.000	
10	0.042	-0.106	127.33	0.000	
11	-0.039	0.001	127.42	0.000	
12	-0.135	-0.142	128.43	0.000	
13	-0.181	0.230	130.33	0.000	
14	-0.157	0.307	131.84	0.000	
15	-0.126	0.257	132.87	0.000	
16	-0.075	0.051	133.25	0.000	

Date: 12/09/13 Time: 13:09 Sample: 1962 2000 Included observations: 39					
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.860	0.860	31.134	0.000	
2	0.740	0.000	54.801	0.000	
3	0.631	-0.023	72.472	0.000	
4	0.538	0.003	85.710	0.000	
5	0.443	-0.063	94.931	0.000	
6	0.354	-0.035	101.02	0.000	
7	0.282	0.001	104.98	0.000	
8	0.208	-0.055	107.22	0.000	
9	0.154	0.020	108.49	0.000	
10	0.114	0.015	109.21	0.000	
11	0.094	0.040	109.72	0.000	
12	0.057	-0.073	109.91	0.000	
13	0.021	-0.036	109.93	0.000	
14	0.017	0.090	109.95	0.000	
15	0.022	0.028	109.99	0.000	
16	0.022	-0.016	110.02	0.000	



Commentaire : les séries « limpt et lpibrt » sont non gaussiennes (prob<5%) et semblent non stationnaires (intégrées d'ordre 1, lag ou nombre de retard = 1).

- Sur Stata, faire :

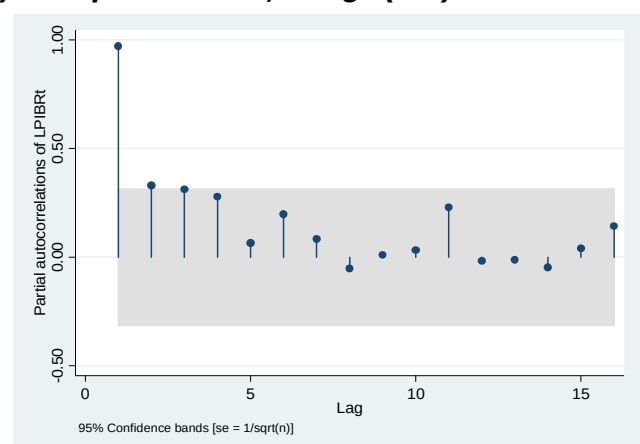
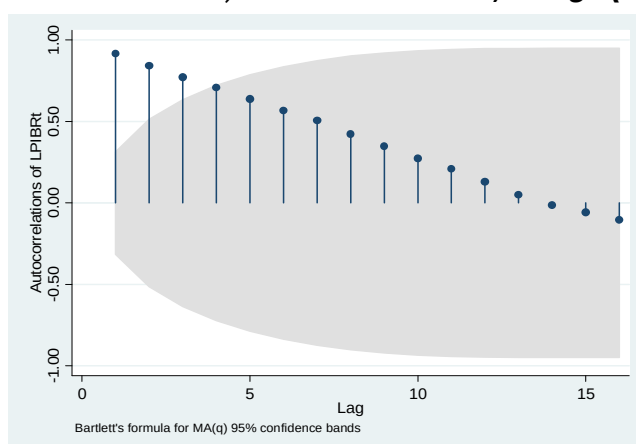
Commande: `corrgram LIMPt, lags(16)`

LAG	AC	PAC	Q	Prob>Q	-1	0	1	-1	0	1
					[Autocorrelation]			[Partial Autocor]		
1	0.8918	0.9387	29.497	0.0000				-----		-----
2	0.8290	0.3057	55.787	0.0000				-----		--
3	0.7379	0.0117	77.286	0.0000				-----		
4	0.6606	0.1933	95.091	0.0000				-----		-
5	0.5841	0.0842	109.49	0.0000				-----		
6	0.4941	-0.1728	120.16	0.0000				---	-	
7	0.4163	0.1044	128.02	0.0000				---		
8	0.2985	-0.1096	132.21	0.0000				--		
9	0.2457	0.2717	135.17	0.0000				-		--
10	0.1468	-0.0713	136.27	0.0000				-		
11	0.0595	-0.1718	136.46	0.0000					-	
12	-0.0504	-0.0104	136.6	0.0000						
13	-0.1226	0.2149	137.47	0.0000						-
14	-0.1818	0.0737	139.5	0.0000			-			
15	-0.2424	0.1801	143.28	0.0000			-			-
16	-0.2550	.	147.7	0.0000		--				

Commande: `corrgram LPIBrt, lags(16)`

LAG	AC	PAC	Q	Prob>Q	-1	0	1	-1	0	1
					[Autocorrelation]			[Partial Autocor]		
1	0.9147	0.9728	35.203	0.0000				-----		-----
2	0.8411	0.3310	65.776	0.0000				-----		--
3	0.7721	0.3133	92.253	0.0000				-----		--
4	0.7072	0.2796	115.1	0.0000				-----		--
5	0.6392	0.0666	134.32	0.0000				-----		
6	0.5671	0.1986	149.9	0.0000				----		-
7	0.5059	0.0850	162.69	0.0000				----		
8	0.4227	-0.0513	171.9	0.0000				---		
9	0.3483	0.0105	178.37	0.0000				--		
10	0.2745	0.0325	182.52	0.0000				--		
11	0.2092	0.2311	185.02	0.0000				-		-
12	0.1303	-0.0157	186.03	0.0000				-		
13	0.0495	-0.0115	186.18	0.0000						
14	-0.0133	-0.0470	186.19	0.0000						
15	-0.0581	0.0411	186.42	0.0000						
16	-0.1036	0.1426	187.16	0.0000						-

Autrement, faire : `ac LIMPt, lags(16)` ou `pac LIMPt, lags(16)`



Avec : `ac` (fonction d'autocorrélation simple) et `pac` (fonction d'autocorrélation partielle).



III.2. Tests de stationnarité (ADF et Phillips-Perron)

a) Test d'Augmented Dickey-fuller (ADF)

Les hypothèses du test sont :

H_0 : la série est non stationnaire $\rightarrow |ADF| < |Mackinnon|$ ou ($prob > 5\%$)

H_1 : la série est stationnaire $\rightarrow |ADF| > |Mackinnon|$ ou ($prob < 5\%$)

► Série : Importation en logarithme (LIMPt)

Commande Stata : `dfuller LIMPt, lags(1) trend regress`

Augmented Dickey-Fuller test for unit root				Number of obs	=	32
				----- Interpolated Dickey-Fuller -----		
	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value		
	Z(t)	-2.686	-4.316	-3.572	-3.223	
MacKinnon approximate p-value for Z(t) = 0.2421						
D.LIMPt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
LIMPt						
L1.	-.5026147	.1871526	-2.69	0.012	-.8859794	-.1192501
LD.	-.090473	.1823063	-0.50	0.624	-.4639105	.2829644
_trend	.0092409	.0035374	2.61	0.014	.0019949	.0164868
_cons	2.824958	1.047789	2.70	0.012	.6786597	4.971257

Commande Eviews : double cliquer sur la série « limpt » pour l'ouvrir, suivre : View/Unit Root test... $\rightarrow$ Choisir (Test type) : ADF, Level, Trend and Intercept (Cfr image ci-dessous) $\rightarrow$ ok (ou faire : **adf** limpt ou encore : **uroot** limpt) :



Null Hypothesis: LIMPT has a unit root Exogenous: Constant, Linear Trend Lag Length: 1 (Fixed)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-2.685589	0.2487
Test critical values:	1% level		-4.273277	
	5% level		-3.557759	
	10% level		-3.212361	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation Dependent Variable: D(LIMPT) Method: Least Squares Date: 12/09/13 Time: 15:17 Sample (adjusted): 1964 1995 Included observations: 32 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LIMPT(-1)	-0.502615	0.187153	-2.685589	0.0120
D(LIMPT(-1))	-0.090473	0.182306	-0.496268	0.6236
C	2.824960	1.047789	2.696114	0.0117
@TREND(1962)	0.009241	0.003537	2.612369	0.0143
R-squared	0.291803	Mean dependent var		0.012842
Adjusted R-squared	0.215925	S.D. dependent var		0.069534
S.E. of regression	0.061571	Akaike info criterion		-2.620780
Sum squared resid	0.106148	Schwarz criterion		-2.437563
Log likelihood	45.93248	F-statistic		3.845670
Durbin-Watson stat	1.980040	Prob(F-statistic)		0.020105

Note : Au seuil de 5%, la série « *limpt* » est non stationnaire en niveau ($|ADF| < |Mackinnon|$ ou $(prob > 5\%)$).

► **Série : PIB réel en logarithme**

Commande Stata : `dfuller LPIBrt, lags(1) trend regress`

Augmented Dickey-Fuller test for unit root			Number of obs = 37		
----- Interpolated Dickey-Fuller -----					
Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value		
Z(t)	-3.459	-4.270	-3.552	-3.211	

MacKinnon approximate p-value for Z(t) = 0.0440					

D.LPIBrt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]

LPIBrt					
L1.	-.7616117	.2201766	-3.46	0.002	-1.209564 -.313659
LD.	.0503153	.1798426	0.28	0.781	-.3155774 .4162079
_trend	.0164457	.0048781	3.37	0.002	.0065211 .0263702
_cons	5.074638	1.457246	3.48	0.001	2.109848 8.039428



Commande Eviews : uroot lpibrt ou ident lpibrt

Null Hypothesis: LPIBRT has a unit root Exogenous: Constant, Linear Trend Lag Length: 1 (Fixed)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-3.459093	0.0590
Test critical values:	1% level		-4.226815	
	5% level		-3.536601	
	10% level		-3.200320	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation Dependent Variable: D(LPIBRT) Method: Least Squares Date: 12/09/13 Time: 15:25 Sample (adjusted): 1964 2000 Included observations: 37 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LPIBRT(-1)	-0.761612	0.220177	-3.459093	0.0015
D(LPIBRT(-1))	0.050315	0.179843	0.279774	0.7814
C	5.074638	1.457247	3.482346	0.0014
@TREND(1962)	0.016446	0.004878	3.371327	0.0019
R-squared	0.353094	Mean dependent var		0.020571
Adjusted R-squared	0.294284	S.D. dependent var		0.043511
S.E. of regression	0.036552	Akaike info criterion		-3.678331
Sum squared resid	0.044091	Schwarz criterion		-3.504178
Log likelihood	72.04912	F-statistic		6.004007
Durbin-Watson stat	1.969188	Prob(F-statistic)		0.002213

Note : Au seuil de 5%, la série « *lpibrt* » est non stationnaire en niveau ($|ADF| < |Mackinnon|$: Cfr résultat Eviews). Toutefois, la série serait stationnaire mais affectée d'une tendance non nuisible (Cfr Stata : $p - value \text{ for } Z(t) < 5\%$)).

b) Test de Philips-Perron (PP)

Les hypothèses du test sont :

H_0 : la série est non stationnaire $\rightarrow |PP| < |Mackinnon|$ ou ($prob > 5\%$)

H_1 : la série est stationnaire $\rightarrow |PP| > |Mackinnon|$ ou ($prob < 5\%$)

► Série : Importation en logarithme (LIMPt)**Commande Stata : pperron LIMPt, lags(1) trend regress**

Phillips-Perron test for unit root		Number of obs =		33	
		Newey-West lags =		1	
	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value	
----- Interpolated Dickey-Fuller -----					
Z(rho)	-15.999	-23.524	-18.508	-15.984	
Z(t)	-3.290	-4.306	-3.568	-3.221	

MacKinnon approximate p-value for Z(t) = 0.0679					
LIMPt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]

LIMPt					
L1.	.4870476	.153194	3.18	0.003	.1741838 .7999114
_trend	.0091574	.0028673	3.19	0.003	.0033016 .0150132
_cons	2.887789	.8605451	3.36	0.002	1.130321 4.645256



Commande Eviews : double cliquer sur la série « limpt » pour l'ouvrir, suivre : View/Unit Root test... → Choisir (Test type): **Philips-Perron**, Level, Trend and Intercept (Cfr image ci-dessous) → ok :

Null Hypothesis: LIMPT has a unit root				
Exogenous: Constant, Linear Trend				
Bandwidth: 2 (Newey-West using Bartlett kernel)				
			Adj. t-Stat	Prob.*
Phillips-Perron test statistic			-3.325118	0.0798
Test critical values:	1% level		-4.262735	
	5% level		-3.552973	
	10% level		-3.209642	
*MacKinnon (1996) one-sided p-values.				
Residual variance (no correction)				0.003308
HAC corrected variance (Bartlett kernel)				0.003203
Phillips-Perron Test Equation				
Dependent Variable: D(LIMPT)				
Method: Least Squares				
Date: 12/09/13 Time: 16:16				
Sample (adjusted): 1963 1995				
Included observations: 33 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LIMPT(-1)	-0.512953	0.153194	-3.348387	0.0022
C	2.887791	0.860545	3.355769	0.0022
@TREND(1962)	0.009157	0.002867	3.193724	0.0033
R-squared	0.273314	Mean dependent var		0.012267
Adjusted R-squared	0.224868	S.D. dependent var		0.068519
S.E. of regression	0.060325	Akaike info criterion		-2.691636
Sum squared resid	0.109173	Schwarz criterion		-2.555590
Log likelihood	47.41200	F-statistic		5.641652
Durbin-Watson stat	2.139247	Prob(F-statistic)		0.008322

Note : la série « limpt » est non stationnaire en niveau ($|PP| < |Mackinnon|$ ou ($prob > 5\%$)).



► Série : PIB réel en logarithme

Commande Stata : `pperron LPIBrt, lags(1) trend regress`

Phillips-Perron test for unit root		Number of obs = 38	
		Newey-West lags = 1	
	Test Statistic	----- Interpolated Dickey-Fuller ----- 1% Critical Value 5% Critical Value 10% Critical Value	
Z(rho)	-27.713	-24.164	-18.888
Z(t)	-4.363	-4.260	-3.548

MacKinnon approximate p-value for Z(t) = 0.0025			

LPIBrt	Coef.	Std. Err.	t P> t [95% Conf. Interval]

LPIBrt			
L1.	.2770148	.1661339	1.67 0.104 - .0602549 .6142845
_trend	.0155874	.0036837	4.23 0.000 .0081091 .0230657
_cons	4.819812	1.101515	4.38 0.000 2.583617 7.056006

Résultat Eviews :

Null Hypothesis: LPIBRT has a unit root Exogenous: Constant, Linear Trend Bandwidth: 1 (Newey-West using Bartlett kernel)				
			Adj. t-Stat	Prob.*
Phillips-Perron test statistic			-4.362969	0.0069
Test critical values:	1% level		-4.219126	
	5% level		-3.533083	
	10% level		-3.198312	
*MacKinnon (1996) one-sided p-values.				
Residual variance (no correction)				0.001163
HAC corrected variance (Bartlett kernel)				0.001178
Phillips-Perron Test Equation Dependent Variable: D(LPIBRT) Method: Least Squares Date: 12/09/13 Time: 16:33 Sample (adjusted): 1963 2000 Included observations: 38 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LPIBRT(-1)	-0.722985	0.166134	-4.351822	0.0001
C	4.819812	1.101515	4.375619	0.0001
@TREND(1962)	0.015587	0.003684	4.231460	0.0002
R-squared	0.354490	Mean dependent var		0.021040
Adjusted R-squared	0.317603	S.D. dependent var		0.043017
S.E. of regression	0.035535	Akaike info criterion		-3.760951
Sum squared resid	0.044195	Schwarz criterion		-3.631668
Log likelihood	74.45808	F-statistic		9.610329
Durbin-Watson stat	1.948598	Prob(F-statistic)		0.000471

Note : Au seuil de 5%, la série « *lpibrt* » est non stationnaire en niveau ($|PP| < |Mackinnon|$: Cfr résultat Eviews). Toutefois, la série serait stationnaire mais affectée d'une tendance non nuisible (Cfr Stata : *p – value for Z(t) < 5%*)).

Constat : les résultats de tests ADF et Philips-Perron convergent.



III.3. Tests de Cointégration (celui de Johansen)

Pour tester l'existence d'une relation de long terme entre des variables, l'on recourt à des procédures statistiques, notamment celle d'Engle et Granger (1987) et celle de Johansen (1988, 1991).

III.3.1. Test d'Engle et Granger (1987)

Le test d'Engle et Granger est souvent sollicité lorsqu'on étudie la relation entre 2 variables. Méthodiquement, suivant Engle et Granger, pour que deux séries soient dites cointégrées :

- o Elles doivent être intégrées de même ordre et de même type ;
- o Le résidu de l'estimation de leur relation de long terme doit être intégré d'ordre « 0 » (stationnaire à niveau, sans trend).

III.3.2. Test de cointégration de Johansen (1988)

Johansen (1988) teste la cointégration à l'aide des estimateurs du maximum de vraisemblance. Il s'agit d'un test de rang de cointégration, utilisé lorsqu'il y a plusieurs vecteurs cointégrant ou dans le cas d'une régression multiple (plus de 2 variables), qui exige que les séries soient intégrées de même ordre⁽¹⁾. Dans ce test, l'on procède par élimination ou exclusion d'hypothèses alternatives pour deux fins : (i) identifier le nombre de relations de cointégration optimal indispensable pour l'estimation d'un vecteur à correction d'erreurs (modèle VECM ou VEC), et (ii) identifier la forme du modèle VECM/VEC en optant pour des équations avec ou sans tendance déterministe, soit des équations avec ou sans tendance linéaire, soit avec ou sans tendance quadratique.

Les différentes formes ou spécifications de modèle VECM en fonction de types de processus sont reprises dans le tableau ci-dessous (tirées de Bourbonnais R., 2015, p.313).

Processus	Formes ou types de spécification VECM				
	I	II	III	IV	V
Tous les processus sont des DS sans dérive	*	*			
Au moins un des processus est un DS avec dérive			*		
Au moins un des processus est un TS				*	
Au moins un processus a une tendance quadratique					*

Les hypothèses du test sont :

H_0 : Pas de relation de cointégration ou rang de cointégration $r = 0 \rightarrow LR < CV$

H_1 : Cointégration ou rang de cointégration $r \geq 1 \rightarrow LR > CV$

¹ Pour des séries intégrées à des ordres différents, recourir au test de cointégration aux bornes.



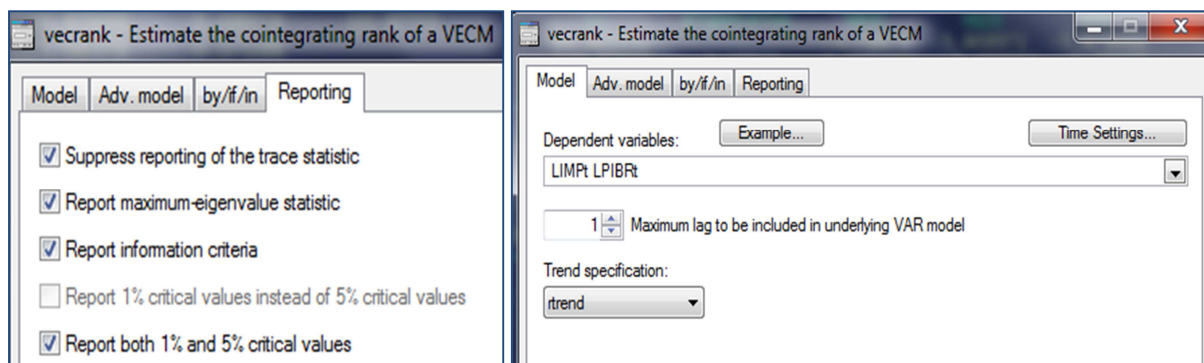
Avec :

- LR : likelihood Ratio (le Ratio de vraisemblance, statistique calculée de Johansen) ;
- CV : Critical value (1%, 5% et 10%).

A noter :

- ♦ Dans l'output du test (Cfr Eviews) :
 - o En ligne, il est indiqué le nombre d'équations/relations de cointégration ou rang de cointégration ;
 - o En colonne, il est indiqué la forme de l'équation de cointégration du modèle VEC.
- ♦ Avec Stata, l'on a la possibilité de comparer deux statistiques calculées (Max-lambda statistics et Trace statistics) aux valeurs critiques de Osterwald-Lenun.

Sur Stata, faire :



Commande : `vecrank LIMPt LPIBrt, trend(rtrend) lags(1) notrace max ic levela`

Johansen tests for cointegration						
Trend: rtrend			Number of obs =		33	
Sample: 1963 1995			Lags =		1	
maximum rank	parms	LL	eigenvalue	max statistic	5% critical value	1% critical value
0	2	99.011551		21.6291	18.96	23.65
1	6	109.82609	0.48078	10.4753	12.52	16.26
2	8	115.06376	0.27199			
maximum rank	parms	LL	eigenvalue	SBIC	HQIC	AIC
0	2	99.011551		-5.78879	-5.848971	-5.879488
1	6	109.82609	0.48078	-6.020398*	-6.20094*	-6.29249
2	8	115.06376	0.27199	-6.125923	-6.366645	-6.488712

Sur Eviews, suivre : Quick/Group Statistics/Cointegration Test → dans List of series, taper : `limpt lpibrt` → ok → dans la boîte de dialogue « Johansen Cointegration Test, cocher « summarize all 5 sets of assumptions » → dans lag intervals, taper : 1 → dans critical value, cocher « osterwald-lenun » → ok : les résultats viennent après cette boîte de dialogue qui résume les dernières étapes de la procédure.



Johansen Cointegration Test

Cointegration Test Specification

Deterministic trend assumption of test

Assume no deterministic trend in data:

☐ 1) No intercept or trend in CE or test VAR

☐ 2) Intercept (no trend) in CE - no intercept in VAR

Allow for linear deterministic trend in data:

☐ 3) Intercept (no trend) in CE and test VAR

☐ 4) Intercept and trend in CE - no trend in VAR

Allow for quadratic deterministic trend in data:

☐ 5) Intercept and trend in CE - linear trend in VAR

Summary:

☒ 6) Summarize all 5 sets of assumptions

* Critical values may not be valid with exogenous variables; do not include C or Trend.

Exog variables*

Lag intervals

1 1

Lag spec for differenced endogenous

Critical Values

☐ MHM

Size 0.05

☒ Osterwald-Lenum

OK Annuler

Date: 12/09/13 Time: 18:29					
Sample: 1962 2000					
Included observations: 32					
Series: LIMPT LPIBRT					
Lags interval: 1 to 1					
Selected (0.05 level*) Number of Cointegrating Relations by Model					
Data Trend:	None	None	Linear	Linear	Quadratic
Test Type	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
Trace	2	0	0	0	2
Max-Eig	0	0	0	0	0
*Critical values based on Osterwald-Lenum (1992)					
Information Criteria by Rank and Model					
Data Trend:	None	None	Linear	Linear	Quadratic
Rank or No. of Ces	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
Log Likelihood by Rank (rows) and Model (columns)					
0	94.75693	94.75693	99.91078	99.91078	99.99511
1	100.3068	100.4114	104.1434	108.1842	108.2258
2	104.0574	104.2999	104.2999	111.9566	111.9566



Akaike Information Criteria by Rank (rows) and Model (columns)					
0	-5.672308	-5.672308	-5.869423	-5.869423	-5.749694
1	-5.769174	-5.713214	-5.883965	-6.074010*	-6.014111
2	-5.753588	-5.643742	-5.643742	-5.997286	-5.997286

Schwarz Criteria by Rank (rows) and Model (columns)					
0	-5.489091	-5.489091	-5.594598*	-5.594598*	-5.383260
1	-5.402740	-5.300976	-5.425922	-5.570163	-5.464460
2	-5.203937	-5.002482	-5.002482	-5.264418	-5.264418

Piste : Il ressort du test de cointégration de Johansen ce qui suit :

- o Il n'existe qu'une seule relation de cointégration suivant le critère d'Akaike qui offre la valeur la plus minimale pour No of CES égal à « 1 » (lecture en ligne) ;
- o La valeur minimale d'Akaike (soit : -6.074010) apparaissant sur la colonne où la caractéristique **Test type** est « *Intercept and trend (Linear)* », l'on déduit la forme de notre modèle à estimer : il s'agit d'une équation linéaire de cointégration avec intercept et tendance.

Sur base de ces informations (surtout la forme du modèle), nous allons reprendre la procédure en spécifiant la forme de notre modèle (soit en précisant les hypothèses que l'on fait sur la présence ou non de l'intercept et du trend dans le VAR ou le VECM retenu) comme suit :

Johansen Cointegration Test

Cointegration Test Specification

Deterministic trend assumption of test

Assume no deterministic trend in data:

☐ 1) No intercept or trend in CE or test VAR

☐ 2) Intercept (no trend) in CE - no intercept in VAR

Allow for linear deterministic trend in data:

☐ 3) Intercept (no trend) in CE and test VAR

☒ 4) Intercept and trend in CE - no trend in VAR

Allow for quadratic deterministic trend in data:

☐ 5) Intercept and trend in CE - linear trend in VAR

Summary:

☐ 6) Summarize all 5 sets of assumptions

* Critical values may not be valid with exogenous variables; do not include C or Trend.

Exog variables*

Lag intervals

11

Lag spec for differenced endogenous

Critical Values

☐ MHM

Size 0.05

☒ Osevald-Lenum

OK Annuler



L'on trouvera dans l'output :

- * Eigen value : valeur propre ;
- * Critical value (CV) : valeur critique ;
- * Likelihood Ratio (LR) : Ratio de vraisemblance ;
- * Rank or No of CES : Rang ou nombre des équations (relations) de cointégration supposées ;
- * Hypothesized : supposition
 - None : aucune ($\text{rang} = 0$)
 - At most 1 : au plus une ($\text{rang} \leq 1$)
 - At most 2 : au plus deux ($\text{rang} \leq 2$)
 - At most 3 : au plus trois ($\text{rang} \leq 3$)

Date: 12/09/13 Time: 20:58 Sample (adjusted): 1964 1995 Included observations: 32 after adjustments Trend assumption: Linear deterministic trend (restricted) Series: LIMPT LPIBRT Lags interval (in first differences): 1 to 1				
Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	5 Percent Critical Value	1 Percent Critical Value
None	0.40374482270438	24.0916132850755	25.32	30.45
At most 1	0.210042680900361	7.54484356597454	12.25	16.26
Trace test indicates no cointegration at both 5% and 1% levels *(**) denotes rejection of the hypothesis at the 5%(1%) level				
Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	5 Percent Critical Value	1 Percent Critical Value
None	0.40374482270438	16.5467697191009	18.96	23.65
At most 1	0.210042680900361	7.54484356597454	12.25	16.26
Max-eigenvalue test indicates no cointegration at both 5% and 1% levels *(**) denotes rejection of the hypothesis at the 5%(1%) level Unrestricted Cointegrating Coefficients (normalized by b*S11*b=I):				
LIMPT	LPIBRT	@TREND(63)		
7.85169853931581	-42.5972651558931	0.822482871575529		
17.3180201328122	-1.15905914413831	-0.294009820696761		
Unrestricted Adjustment Coefficients (alpha):				
D(LIMPT)	-0.00136037031660071	-0.0292370094359988		
D(LPIBRT)	0.025338531007357	-0.00765105857387352		
1 Cointegrating Equation(s):			Log likelihood	108.1841 60207916
Normalized cointegrating coefficients (standard error in parentheses)				
LIMPT	LPIBRT	@TREND(63)		



1	-5.42522932364199	0.104752222395843
	1.1212310856303	0.0257596908460973
Adjustment coefficients (standard error in parentheses)		
D(LIMPT)	-0.0106812176277825	0.0946911846410738
D(LPIBRT)	0.198950506898872	0.0519737943800937

Résultats : Il ressort des résultats du test de cointégration, tenant compte de la spécification VECM retenue, qu'il n'existe aucune relation de cointégration entre les variables considérées. En fait, les statistiques calculées de la trace (*Trace Statistic*) et la valeur propre maximale (*Max-Eigen Statistic*) sont inférieures aux valeurs critiques à 5%, voir même 1%, ce qui amène à accepter l'hypothèse nulle selon laquelle le rang ou le nombre de vecteurs cointégrants est nul ($r=0$ ou « Non »).

Hypothèses		Trace		Valeur propre maximale	
H_0	H_1	Statistique	Val. Critique 5%	statistique	Val. Critique 5%
$r = 0$	$r \leq 1$	24.09	25.32	16.54	18.96

III.4. Estimation d'un MCE

L'on voudrait ici tester la cointégration entre nos variables et estimer éventuellement un modèle à correction d'erreur/MCE par l'approche proposée par Engle et Granger (1987) dont les étapes sont (Bourbonnais R., 2015, p. 308) :

- (i) Estimer, par les MCO, la relation de long terme et calculer les résidus de l'estimation (les variables sont prises en logarithme) :

$$\lim p_t = a_0 + a_1 \ln pibr_t + e_t \dots \dots [3.1]$$

$$\hat{e}_t = \lim p_t - \hat{a}_0 - \hat{a}_1 \ln pibr_t \dots \dots [3.2]$$

A ce niveau, pour vérifier l'existence d'une relation de long terme entre nos deux variables, soit tester la cointégration, l'on doit tester la stationnarité des résidus estimés. Les résidus, exprimés par l'équation 3.2, doivent être stationnaires en niveau sans tendance (ni constante) pour valider l'hypothèse d'une cointégration entre les variables étudiées.

- (ii) Estimer par les MCO le modèle à correction d'erreur, il intègre dans la dynamique de long terme les ajustements ou fluctuations de court terme, spécifié comme suit :

$$\Delta \lim p_t = b_1 \Delta \ln pibr_t + \lambda e_{t-1} + u_t \dots \dots [3.3]$$

Avec « λ » la force de rappel vers l'équilibre ou coefficient d'ajustement. Pour confirmer l'hypothèse de cointégration entre nos deux séries, les conditions suivantes doivent être vérifiées :



- $0 < |\lambda| < 1$: condition de stabilité du processus.
- « λ » est négatif pour garantir l'ajustement ou convergence vers l'équilibre.
- « λ » est statistiquement significatif.

Cette démarche d'Engle et Granger va nous aider à confirmer ou rejeter les résultats précédents fournis par le test de cointégration de Johansen (1988).

a) Estimation de la relation de long terme

Sur Eviews, taper : **ls log(imp) c log(pibr)**

Dependent Variable: LOG(IMP) Method: Least Squares Date: 01/13/18 Time: 16:13 Sample (adjusted): 1962 1995 Included observations: 34 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.648755	0.346992	1.869655	0.0707
LOG(PIBR)	0.750630	0.049422	15.18820	0.0000
R-squared	0.878179	Mean dependent var		5.916234
Adjusted R-squared	0.874373	S.D. dependent var		0.182731
S.E. of regression	0.064767	Akaike info criterion		-2.579013
Sum squared resid	0.134233	Schwarz criterion		-2.489227
Log likelihood	45.84322	Hannan-Quinn criter.		-2.548393
F-statistic	230.6815	Durbin-Watson stat		1.064827
Prob(F-statistic)	0.000000			

b) Test de stationnarité sur les résidus estimés

Sur Eviews, taper : **genr e=resid**
adf e

Null Hypothesis: E has a unit root				
Exogenous: None				
Lag Length: 0 (Automatic based on SIC, MAXLAG=8)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-3.798129	0.0004
Test critical values:			1% level	-2.636901
			5% level	-1.951332
			10% level	-1.610747
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(E)				
Method: Least Squares				
Date: 01/13/18 Time: 16:15				
Sample (adjusted): 1963 1995				
Included observations: 33 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.



E(-1)	-0.578999	0.152443	-3.798129	0.0006
-------	-----------	----------	-----------	--------

Comme l'on peut le voir, $|ADF| > |McKinnon|$, les résidus issus de la relation de long terme sont stationnaires en niveau sans tendance ni constante. Ceci présage l'existence d'une relation de cointégration entre nos deux séries.

c) Estimation du MCE

Sur Eviews, taper : **ls d(log(limp)) d(log(lpibr)) e(-1)**

Dependent Variable: D(LOG(LIMP)) Method: Least Squares Date: 01/13/18 Time: 16:41 Sample (adjusted): 1963 1995 Included observations: 33 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LOG(LPIBR))	0.619282	0.223128	2.775457	0.0093
E(-1)	-0.096438	0.025565	-3.772335	0.0007
R-squared	0.381699	Mean dependent var		0.002065
Adjusted R-squared	0.361754	S.D. dependent var		0.011627
S.E. of regression	0.009289	Akaike info criterion		-6.461317
Sum squared resid	0.002675	Schwarz criterion		-6.370620
Log likelihood	108.6117	Hannan-Quinn criter.		-6.430800
Durbin-Watson stat	2.128321			

Les résultats d'estimations sont résumés comme suit :

$$\Delta \text{limp}_t = 0,61 * \Delta \text{lpibr}_t - \mathbf{0,09} * \mathbf{e_{t-1}}$$

(2,77) (-3,77)

$$R^2 = 0,38 ; DW = 2,12 ; N = 2,12 ; (.) = t - \text{stat}$$

Note :

- La force de rappel est négative, comprise entre 0 et 1 en valeur absolue, et est statistiquement significative. L'on en déduit que nos deux séries sont cointégrées et que l'on peut estimer un modèle à correction d'erreur, contrairement aux résultats trouvés précédemment (Cfr résultats du test de cointégration Johansen).
- Tous les paramètres sont statistiquement significatifs (t-stat > 2), mais le modèle spécifié explique seulement 38% des variations des importations (le modèle a passé avec succès tous les tests post-estimations, sauf l'hypothèse de stabilité qui est violée). La prise en compte d'autres variables explicatives pertinentes serait de nature à améliorer le pouvoir explicatif du modèle.
- **Force de rappel** : disons que les chocs sur les importations (imp) se corrigent à environ 9% par l'effet de « feed back » ; autrement dit, l'on arrive à ajuster 9% du déséquilibre entre le niveau désiré et effectif des importations.



- **Retard moyen** = $|1/0,096|$: un choc constaté sur les importations est entièrement résorbé au bout de 10 ans et 5 mois en moyenne (l'ajustement est lent).
- **Elasticité à court terme** : à court terme, si le PIB réel augmente de 1%, les importations varie, dans le même sens, de 0,6%.

B. ESTIMATION D'UN VECM

Cas pratique 3.2 : Vérification de la cointégration et estimation d'un vecteur à correction d'erreur/VECM (taux de change, masse monétaire et inflation).

Les variables considérées, prises en logarithme sur une fréquence mensuelle, sont observées sur la RDC et s'étale sur la période 2002-2016. Les séries sont : le taux de change indicatif (ltxchange), l'indice des prix à la consommation (lipc) et la masse monétaire au sens large (lmm).

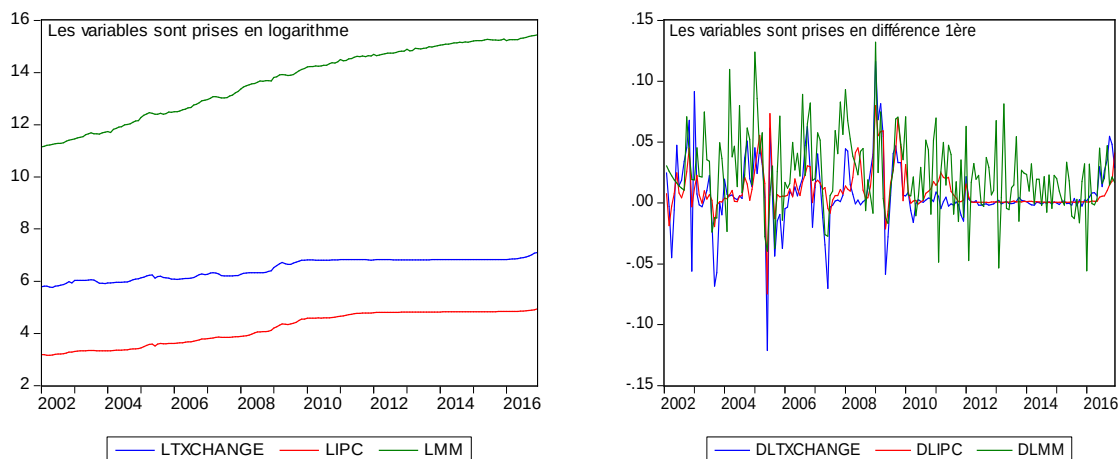
Les résultats de test de cointégration (test de Johansen (1988, 1991)), si l'on arrive à valider l'hypothèse de cointégration entre variables, peuvent renseigner 1 ou plusieurs vecteurs cointégrants, ce qui va conduire à estimer un MCE ou VECM selon le cas. Le VECM, rappelons-le, est estimé lorsqu'on a plus de 2 variables intégrées de même ordre (ordre 1).

Ci-dessous une synthèse d'étapes à suivre pour estimer un modèle vectoriel à correction d'erreur :

- Tester la stationnarité des séries (graphique, ADF).
- Déterminer le lag ou décalage optimal du modèle VAR associé.
- Tester l'hypothèse de cointégration entre variables par le test de cointégration de Johansen (test de la Trace et celui de la valeur propre maximale).
- Si cointégration il y a, identifier le nombre de vecteurs cointégrants ou nombre de relations de cointégration et la spécification ou forme du VECM à estimer.
- Estimer le VECM par le maximum de vraisemblance et valider les résultats (tests usuels : absence d'autocorrélation et hétéroscédasticité, exogénéité faible, etc.).
- Valider le VECM estimé.

Nous nous abstenons de reprendre certaines commandes Eviews ou Stata parce que appliquées plus haut.



a) Graphique**b) Stationnarité des variables**

Variables	Phillips-Perron	ADF en Niveau	McKinnon à 5%	ADF en différence	Stationnarité
Lmm	-2.62	-2.89	-2.87	-12.66	I(1)
Lipc	-	-1.30	-2.87	-8.55	I(1)
Ltxchange	-	2.38	-2.87	-9.40	I(1)

L'on trouve que toutes nos variables, au départ non stationnaires en niveau, deviennent stationnaires après la 1^{ère} différence. Elles sont intégrées de même ordre, ce qui suppose l'existence d'une relation de cointégration entre elles.

Ci-dessous, les différents résultats de test de stationnarité sur nos séries.

Null Hypothesis: LMM has a unit root				
Exogenous: Constant				
Lag Length: 0 (Automatic based on SIC, MAXLAG=13)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-2.892702	0.0482
Test critical values:			1% level	-3.466994
			5% level	-2.877544
			10% level	-2.575381
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(LMM)				
Method: Least Squares				
Date: 01/10/18 Time: 17:41				
Sample (adjusted): 2002M02 2016M12				
Included observations: 179 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LMM(-1)	-0.005015	0.001734	-2.892702	0.0043
C	0.092385	0.023768	3.886991	0.0001



Null Hypothesis: LMM has a unit root				
Exogenous: Constant				
Bandwidth: 4 (Newey-West using Bartlett kernel)				
			Adj. t-Stat	Prob.*
Phillips-Perron test statistic			-2.626626	0.0895
Test critical values:	1% level	-3.466994		
	5% level	-2.877544		
	10% level	-2.575381		
*MacKinnon (1996) one-sided p-values.				
Residual variance (no correction)			0.000963	
HAC corrected variance (Bartlett kernel)			0.001199	
Phillips-Perron Test Equation				
Dependent Variable: D(LMM)				
Method: Least Squares				
Date: 01/10/18 Time: 17:40				
Sample (adjusted): 2002M02 2016M12				
Included observations: 179 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LMM(-1)	-0.005015	0.001734	-2.892702	0.0043
C	0.092385	0.023768	3.886991	0.0001

Null Hypothesis: DLMM has a unit root				
Exogenous: Constant, Linear Trend				
Bandwidth: 3 (Newey-West using Bartlett kernel)				
			Adj. t-Stat	Prob.*
Phillips-Perron test statistic			-12.66318	0.0000
Test critical values:	1% level	-4.010440		
	5% level	-3.435269		
	10% level	-3.141649		
*MacKinnon (1996) one-sided p-values.				
Residual variance (no correction)				0.000964
HAC corrected variance (Bartlett kernel)				0.001061
Phillips-Perron Test Equation				
Dependent Variable: D(DLMM)				
Method: Least Squares				
Date: 01/10/18 Time: 17:42				
Sample (adjusted): 2002M03 2016M12				
Included observations: 178 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLMM(-1)	-0.952469	0.075506	-12.61445	0.0000
C	0.034422	0.005483	6.278001	0.0000
@TREND(2002M01)	-0.000129	4.68E-05	-2.748090	0.0066



Null Hypothesis: DLTXCHANGE has a unit root				
Exogenous: Constant				
Lag Length: 0 (Automatic based on SIC, MAXLAG=13)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-9.409031	0.0000
Test critical values:			1% level	-3.467205
			5% level	-2.877636
			10% level	-2.575430
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(DLTXCHANGE)				
Method: Least Squares				
Date: 01/10/18 Time: 17:30				
Sample (adjusted): 2002M03 2016M12				
Included observations: 178 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLTXCHANGE(-1)	-0.669972	0.071205	-9.409031	0.0000
C	0.004836	0.001983	2.438757	0.0157

Null Hypothesis: DLIPC has a unit root				
Exogenous: Constant				
Lag Length: 0 (Automatic based on SIC, MAXLAG=13)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-8.556092	0.0000
Test critical values:			1% level	-3.467205
			5% level	-2.877636
			10% level	-2.575430
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(DLIPC)				
Method: Least Squares				
Date: 01/10/18 Time: 17:26				
Sample (adjusted): 2002M03 2016M12				
Included observations: 178 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLIPC(-1)	-0.596725	0.069743	-8.556092	0.0000
C	0.005976	0.001372	4.356147	0.0000

c) *Décalage optimal*

Pour déterminer le décalage optimal, estimer d'abord le VAR. Dans Eviews, suivre (Cfr menu des raccourcis) : *Quick/Estimate VAR..* → remplir comme ci-dessous dans la boîte de dialogue qui apparaît :



* Ensuite, cliquer sur « Ok ».

* dans l'output de résultat du VAR estimé, suivre : **View/Lag Length Criteria...** → **ok**. L'on a le résultat ci-bas.

VAR Lag Order Selection Criteria Endogenous variables: LIPC LTXCHANGE LMM Exogenous variables: C Date: 01/10/18 Time: 17:45 Sample: 2002M01 2016M12 Included observations: 172						
Lag	LogL	LR	FPE	AIC	SC	HQ
0	130.8195	NA	4.54e-05	-1.486274	-1.431376	-1.464000
1	1293.135	2270.570	6.81e-11	-14.89692	-14.67733	-14.80783
2	1320.935	53.33690*	5.47e-11*	-15.11553*	-14.73124*	-14.95961*
3	1324.203	6.156199	5.85e-11	-15.04888	-14.49989	-14.82614
4	1327.345	5.808373	6.26e-11	-14.98075	-14.26708	-14.69120
5	1334.719	13.37641	6.39e-11	-14.96185	-14.08348	-14.60547
6	1343.441	15.51693	6.42e-11	-14.95862	-13.91555	-14.53542
7	1351.299	13.70578	6.51e-11	-14.94534	-13.73758	-14.45532
8	1359.280	13.64188	6.60e-11	-14.93349	-13.56103	-14.37665

* indicates lag order selected by the criterion
 LR: sequential modified LR test statistic (each test at 5% level)
 FPE: Final prediction error
 AIC: Akaike information criterion
 SC: Schwarz information criterion
 HQ: Hannan-Quinn information criterion

d) Test de l'hypothèse de cointégration

Pour tester l'hypothèse de cointégration entre nos séries, l'on recourt au test de cointégration de Johansen (1988) dont on rappelle les hypothèses :

H_0 : Pas de relation de cointégration ou rang de cointégration $r = 0 \rightarrow LR < CV$

H_1 : Cointégration ou rang de cointégration $r \geq 1 \rightarrow LR > CV$

Avec :

- LR : likelihood Ratio (le Ratio de vraisemblance, statistique calculée de Johansen) ;
- CV : Critical value (1%, 5% et 10%).



Date: 01/10/18 Time: 17:46 Sample: 2002M01 2016M12 Included observations: 177 Series: LMM LIPC LTXCHANGE Lags interval: 1 to 2					
Selected (0.05 level*) Number of Cointegrating Relations by Model					
Data Trend:	None	None	Linear	Linear	Quadratic
Test Type	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
Trace	1	1	1	1	1
Max-Eig	1	2	0	1	1
*Critical values based on MacKinnon-Haug-Michelis (1999)					
Information Criteria by Rank and Model					
Data Trend:	None	None	Linear	Linear	Quadratic
Rank or No. of CEs	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
Log Likelihood by Rank (rows) and Model (columns)					
0	1324.445	1324.445	1343.185	1343.185	1349.861
1	1348.914	1349.060	1352.167	1361.125	1365.256
2	1350.946	1357.110	1357.117	1366.311	1367.845
3	1350.996	1358.867	1358.867	1368.218	1368.218
Akaike Information Criteria by Rank (rows) and Model (columns)					
0	-14.76209	-14.76209	-14.93995	-14.93995	-14.98148
1	-14.97078	-14.96113	-14.97364	-15.06356	-15.08764*
2	-14.92594	-14.97300	-14.96177	-15.04306	-15.04910
3	-14.85871	-14.91375	-14.91375	-14.98552	-14.98552
Schwarz Criteria by Rank (rows) and Model (columns)					
0	-14.43909	-14.43909	-14.56312*	-14.56312*	-14.55082
1	-14.54012	-14.51252	-14.48914	-14.56112	-14.54931
2	-14.38761	-14.39878	-14.36961	-14.41501	-14.40310
3	-14.21271	-14.21392	-14.21392	-14.23186	-14.23186

Les résultats du test ci-dessus renseignent, au regard de la valeur AIC la plus minimale (AIC = -15,08), 1 vecteur cointégrant pour un VECM avec tendance quadratique, soit la 5^{ème} spécification parmi les cinq disponibles.

Processus	Formes ou types de spécification VECM				
	I	II	III	IV	V
Tous les processus sont des DS sans dérive	*	*			
Au moins un des processus est un DS avec dérive			*		
Au moins un des processus est un TS				*	
Au moins un processus a une tendance quadratique					*

En reprenant la procédure du test, tenant compte de la spécification VECM appropriée aux séries, l'on obtient les résultats ci-dessous.



Date: 01/10/18 Time: 17:49

Sample (adjusted): 2002M04 2016M12

Included observations: 177 after adjustments

Trend assumption: Quadratic deterministic trend

Series: LMM LIPC LTXCHANGE

Lags interval (in first differences): 1 to 2

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.159661	36.71431	35.01090	0.0325
At most 1	0.028837	5.925159	18.39771	0.8746
At most 2	0.004206	0.746014	3.841466	0.3877

Trace test indicates 1 cointegrating eqn(s) at the 0.05 level

* denotes rejection of the hypothesis at the 0.05 level

**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.159661	30.78915	24.25202	0.0059
At most 1	0.028837	5.179145	17.14769	0.8889
At most 2	0.004206	0.746014	3.841466	0.3877

Max-eigenvalue test indicates 1 cointegrating eqn(s) at the 0.05 level

* denotes rejection of the hypothesis at the 0.05 level

**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegrating Coefficients (normalized by b*S11*b=I):

LMM	LIPC	LTXCHANGE
7.261092	-15.79045	13.57512
5.827733	-1.483937	-11.48240
1.188306	11.66726	-10.78858

Unrestricted Adjustment Coefficients (alpha):

D(LMM)	0.001273	0.003711	0.001308
D(LIPC)	0.003711	0.001927	-0.000310
D(LTXCHANGE)	-0.000809	0.003976	-0.000579

1 Cointegrating Equation(s):

Log likelihood

1365.256

Normalized cointegrating coefficients (standard error in parentheses)

LMM	LIPC	LTXCHANGE
1.000000	-2.174666	1.869571
	(0.38675)	(0.50460)

Adjustment coefficients (standard error in parentheses)

D(LMM)	0.009245
	(0.01674)
D(LIPC)	0.026945
	(0.00839)
D(LTXCHANGE)	-0.005874
	(0.01408)

2 Cointegrating Equation(s):

Log likelihood

1367.845



Normalized cointegrating coefficients (standard error in parentheses)

LMM	LIPC	LTXCHANGE
1.000000	0.000000	-2.479545 (0.73597)
0.000000	1.000000	-1.999901 (0.35481)

Adjustment coefficients (standard error in parentheses)

D(LMM)	0.030874 (0.02130)	-0.025612 (0.03628)
D(LIPC)	0.038174 (0.01066)	-0.061456 (0.01816)
D(LTXCHANGE)	0.017294 (0.01782)	0.006874 (0.03036)

Les statistiques calculées de la trace et la valeur propre maximale sont inférieures aux valeurs critiques (seuil de 5%) pour le rang de cointégration égale à 1, ce qui traduit l'existence d'un vecteur cointégrant. Ainsi, nous pouvons estimer un MCE ou un VECM.

e) Estimation d'un VECM

Pour estimer un VECM, sur EViews, suivre la même procédure pour estimer un VAR : Quick/Estimate VAR..→ remplir comme ci-dessous dans la boîte de dialogue qui apparaît→ cliquer sur « Ok ».

Vector Error Correction Estimates

Date: 01/10/18 Time: 17:54
 Sample (adjusted): 2002M04 2016M12
 Included observations: 177 after adjustments
 Standard errors in () & t-statistics in []

Cointegrating Eq:	CointEq1
LIPC(-1)	1.000000
LTXCHANGE(-1)	-0.859705 (0.14315) [-6.00574]
LMM(-1)	-0.459841 (0.08478)

Equation de Long terme



	[-5.42401]		
@TREND(02M01)	0.006085		
C	7.112307		
Error Correction:	D(LIPC)	D(LTXCHANGE)	D(LMM)
CointEq1	-0.058597 (0.01824) [-3.21331]	0.012774 (0.03061) [0.41730]	-0.020105 (0.03640) [-0.55229]
D(LIPC(-1))	0.184047 (0.11737) [1.56803]	-0.086069 (0.19702) [-0.43685]	0.138672 (0.23431) [0.59184]
D(LIPC(-2))	-0.091438 (0.11587) [-0.78914]	-0.133231 (0.19450) [-0.68500]	-0.389500 (0.23131) [-1.68392]
D(LTXCHANGE(-1))	0.052045 (0.07129) [0.73008]	0.298390 (0.11966) [2.49367]	0.254704 (0.14230) [1.78986]
D(LTXCHANGE(-2))	0.075024 (0.07092) [1.05784]	0.116895 (0.11905) [0.98192]	0.127859 (0.14158) [0.90311]
D(LMM(-1))	0.038536 (0.04343) [0.88739]	0.115142 (0.07289) [1.57961]	-0.089413 (0.08669) [-1.03144]
D(LMM(-2))	-0.042256 (0.04350) [-0.97137]	0.006571 (0.07302) [0.08999]	0.003984 (0.08684) [0.04588]
C	0.012823 (0.00352) [3.63892]	0.003741 (0.00592) [0.63237]	0.039845 (0.00703) [5.66413]
@TREND(02M01)	-4.92E-05 (2.5E-05) [-1.97261]	-2.18E-06 (4.2E-05) [-0.05210]	-0.000156 (5.0E-05) [-3.12550]
R-squared	0.241325	0.133454	0.124319
Adj. R-squared	0.205197	0.092190	0.082620
Sum sq. resids	0.039659	0.111742	0.158038
S.E. equation	0.015364	0.025790	0.030671
F-statistic	6.679824	3.234147	2.981326
Log likelihood	492.5655	400.8904	370.2127
Akaike AIC	-5.464017	-4.428141	-4.081499
Schwarz SC	-5.302518	-4.266641	-3.920000
Mean dependent	0.010049	0.007264	0.023915
S.D. dependent	0.017234	0.027068	0.032022
Determinant resid covariance (dof adj.)	4.69E-11		
Determinant resid covariance	4.01E-11		
Log likelihood	1365.256		
Akaike information criterion	-15.08764		
Schwarz criterion	-14.54931		

Force de rappel

Variations de court terme



Dans l'output de l'estimation, suivre : View/Representations : Afficher la spécification du modèle et les équations estimées.

```

Estimation Proc:
=====
EC(E,1) 1 2 LIPC LTXCHANGE LMM

VAR Model:
=====
D(LIPC) = A(1,1)*(B(1,1)*LIPC(-1) + B(1,2)*LTXCHANGE(-1) + B(1,3)*LMM(-1) + B(1,4)*@TREND(02M01) +
B(1,5)) + C(1,1)*D(LIPC(-1)) + C(1,2)*D(LIPC(-2)) + C(1,3)*D(LTXCHANGE(-1)) + C(1,4)*D(LTXCHANGE(-2)) +
C(1,5)*D(LMM(-1)) + C(1,6)*D(LMM(-2)) + C(1,7) + C(1,8)*@TREND(02M01)

D(LTXCHANGE) = A(2,1)*(B(1,1)*LIPC(-1) + B(1,2)*LTXCHANGE(-1) + B(1,3)*LMM(-1) +
B(1,4)*@TREND(02M01) + B(1,5)) + C(2,1)*D(LIPC(-1)) + C(2,2)*D(LIPC(-2)) + C(2,3)*D(LTXCHANGE(-1)) +
C(2,4)*D(LTXCHANGE(-2)) + C(2,5)*D(LMM(-1)) + C(2,6)*D(LMM(-2)) + C(2,7) + C(2,8)*@TREND(02M01)

D(LMM) = A(3,1)*(B(1,1)*LIPC(-1) + B(1,2)*LTXCHANGE(-1) + B(1,3)*LMM(-1) + B(1,4)*@TREND(02M01) +
B(1,5)) + C(3,1)*D(LIPC(-1)) + C(3,2)*D(LIPC(-2)) + C(3,3)*D(LTXCHANGE(-1)) + C(3,4)*D(LTXCHANGE(-2)) +
C(3,5)*D(LMM(-1)) + C(3,6)*D(LMM(-2)) + C(3,7) + C(3,8)*@TREND(02M01)

VAR Model - Substituted Coefficients:
=====
D(LIPC) = - 0.0585970917448*( LIPC(-1) - 0.859704744192*LTXCHANGE(-1) - 0.459840767851*LMM(-1) +
0.00608541290206*@TREND(02M01) + 7.11230679623 ) + 0.184046599259*D(LIPC(-1)) -
0.0914381924321*D(LIPC(-2)) + 0.0520447385675*D(LTXCHANGE(-1)) + 0.0750237080815*D(LTXCHANGE(-
2)) + 0.0385356647808*D(LMM(-1)) - 0.0422564057558*D(LMM(-2)) + 0.0128234172025 - 4.92074274615e-
05*@TREND(02M01)

D(LTXCHANGE) = 0.012773588715*( LIPC(-1) - 0.859704744192*LTXCHANGE(-1) - 0.459840767851*LMM(-1)
+ 0.00608541290206*@TREND(02M01) + 7.11230679623 ) - 0.0860685878214*D(LIPC(-1)) -
0.133231410976*D(LIPC(-2)) + 0.298389572005*D(LTXCHANGE(-1)) + 0.116894627793*D(LTXCHANGE(-2)) +
0.115142304677*D(LMM(-1)) + 0.00657133242735*D(LMM(-2)) + 0.00374060602766 - 2.18148932355e-
06*@TREND(02M01)

D(LMM) = - 0.0201046944543*( LIPC(-1) - 0.859704744192*LTXCHANGE(-1) - 0.459840767851*LMM(-1) +
0.00608541290206*@TREND(02M01) + 7.11230679623 ) + 0.138672166502*D(LIPC(-1)) -
0.389500340203*D(LIPC(-2)) + 0.2547036735*D(LTXCHANGE(-1)) + 0.127859136453*D(LTXCHANGE(-2)) -
0.089413223092*D(LMM(-1)) + 0.00398428535914*D(LMM(-2)) + 0.0398449922365 -
0.000155639275223*@TREND(02M01)

```

f) Validation du VECM estimé

De résultats de l'estimation du VECM, l'on note que les coefficients de termes de force de rappel des variables « *lipc* et *lmm* » ont les signes attendus (négatifs), par contre celui de la variable « *ltxchange* » n'a pas le signe attendu. En outre, tous les coefficients de la « relation de long terme » sont statistiquement significatifs ; par contre, seule la force de rappel associée à l'équation de la variable « *lipc* » est statistiquement significative (le R^2 du VECM estimé est aussi faible, ce qui traduit l'omission de certaines variables explicatives pertinentes). Il en ressort que les variables « *ltxchange* » et « *lmm* », statistiquement non significatives, sont « faiblement exogènes ». Pour en être convaincu, l'on peut procéder par un « test de contrainte » ou « test d'exogénéité faible » sur les coefficients de nos trois variables sous étude. Pour ce faire, estimer le VECM en précisant des restrictions [$A(1,1)=0$; $A(2,1)=0$; $A(3,1)=0$] dans la boîte de dialogue ci-dessous (Cfr boîte de dialogue pour l'estimation d'un VECM).



Restrictions may be placed on the coefficients $B(r,k)$ of the r -th cointegrating relation:

$$B(r,1)*LIPC + B(r,2)*LTXCHANGE + B(r,3)*LMM$$

VEC Coefficient Restrictions

☒ Impose Restrictions

Enter restriction: (Example: $B(1,1)=1, A(2,1)=0$)

$A(3,1)=0$

Optimization

Max Iterations: 500

Convergence: 0.0001

OK Annuler

Les hypothèses du test (statistique de chi-deux) sont :

$H_0 : A(i, j) = 0, (prob > 5\%) : \text{la variable considérée est faiblement exogène}$

$H_1 : A(i, j) \neq 0, (prob < 5\%) : \text{la variable considérée n'est pas faiblement exogène}$

Les résultats du test sont résumés comme suit :

Cointegration Restrictions			
	$A(1,1)=0$	$A(2,1)=0$	$A(3,1)=0$
Chi-square(1)	9,044	0,157	0,289
Probability	0,002	0,691	0,590
Décision	rejet H_0	accepte H_0	accepte H_0
Not all cointegrating vectors are identified			
LR test for binding restrictions (rank = 1):			

Vector Error Correction Estimates	
Date: 01/11/18 Time: 08:12	
Sample (adjusted): 2002M04 2016M12	
Included observations: 177 after adjustments	
Standard errors in () & t-statistics in []	
Cointegration Restrictions:	
$A(1,1)=0$	
Convergence achieved after 5 iterations.	
Not all cointegrating vectors are identified	
LR test for binding restrictions (rank = 1):	
Chi-square(1)	9.044638
Probability	0.002635

Vector Error Correction Estimates	
Date: 01/11/18 Time: 08:15	
Sample (adjusted): 2002M04 2016M12	
Included observations: 177 after adjustments	
Standard errors in () & t-statistics in []	
Cointegration Restrictions:	
$A(2,1)=0$	
Convergence achieved after 3 iterations.	
Not all cointegrating vectors are identified	



LR test for binding restrictions (rank = 1):	
Chi-square(1)	0.157447
Probability	0.691518

Vector Error Correction Estimates	
Date: 01/11/18 Time: 08:18	
Sample (adjusted): 2002M04 2016M12	
Included observations: 177 after adjustments	
Standard errors in () & t-statistics in []	

Cointegration Restrictions:	
A(3,1)=0	
Convergence achieved after 4 iterations.	
Not all cointegrating vectors are identified	
LR test for binding restrictions (rank = 1):	
Chi-square(1)	0.289074
Probability	0.590814

Les résultats des tests ci-dessus confirment l'exogénéité faible des variables « *ltxchange* » et « *lmm* ». A présent, tenant compte de ces résultats, nous estimons à nouveau un VECM *partiel ou contraint*. Ci-dessous la procédure (boîte de dialogue) et les résultats du VECM partiel estimé.

Vector Error Correction Estimates	
Date: 01/10/18 Time: 19:37	
Sample (adjusted): 2002M04 2016M12	
Included observations: 177 after adjustments	
Standard errors in () & t-statistics in []	

Cointegration Restrictions:	
A(2,1)=0, A(3,1)=0	
Convergence achieved after 3 iterations.	
Not all cointegrating vectors are identified	
LR test for binding restrictions (rank = 1) :	
Chi-square(2)	0.859174
Probability	0.650778

Cointegrating Eq:	CointEq1
-------------------	----------

LIPC(-1)	-16.08326
LTXCHANGE(-1)	13.75844



LMM(-1)	7.270948		
@TREND(02M01)	-0.094216		
C	-112.5721		
Error Correction:	D(LIPC)	D(LTXCHANGE)	D(LMM)
CointEq1	0.004027 (0.00072) [5.57658]	0.000000 (0.00000) [NA]	0.000000 (0.00000) [NA]
D(LIPC(-1))	0.184844 (0.11723) [1.57682]	-0.087482 (0.19684) [-0.44442]	0.139318 (0.23409) [0.59515]
D(LIPC(-2))	-0.090379 (0.11571) [-0.78111]	-0.134655 (0.19429) [-0.69305]	-0.388779 (0.23105) [-1.68265]
D(LTXCHANGE(-1))	0.051056 (0.07122) [0.71682]	0.298910 (0.11960) [2.49924]	0.254273 (0.14223) [1.78778]
D(LTXCHANGE(-2))	0.073889 (0.07087) [1.04255]	0.117326 (0.11901) [0.98585]	0.127415 (0.14153) [0.90029]
D(LMM(-1))	0.038448 (0.04341) [0.88568]	0.114875 (0.07289) [1.57593]	-0.089358 (0.08669) [-1.03083]
D(LMM(-2))	-0.042376 (0.04349) [-0.97441]	0.006344 (0.07303) [0.08687]	0.004019 (0.08684) [0.04628]
C	0.012822 (0.00352) [3.64257]	0.003792 (0.00591) [0.64159]	0.039829 (0.00703) [5.66610]
@TREND(02M01)	-4.92E-05 (2.5E-05) [-1.97258]	-2.39E-06 (4.2E-05) [-0.05716]	-0.000156 (5.0E-05) [-3.12498]
R-squared	0.241816	0.133364	0.124301
Adj. R-squared	0.205712	0.092096	0.082601
Sum sq. resid	0.039633	0.111754	0.158041
S.E. equation	0.015359	0.025791	0.030671
F-statistic	6.697752	3.231638	2.980847
Log likelihood	492.6228	400.8813	370.2109
Akaike AIC	-5.464665	-4.428037	-4.081479
Schwarz SC	-5.303166	-4.266538	-3.919980
Mean dependent	0.010049	0.007264	0.023915
S.D. dependent	0.017234	0.027068	0.032022
Determinant resid covariance (dof adj.)		4.69E-11	
Determinant resid covariance		4.01E-11	
Log likelihood		1364.826	
Akaike information criterion		-15.08278	
Schwarz criterion		-14.54445	

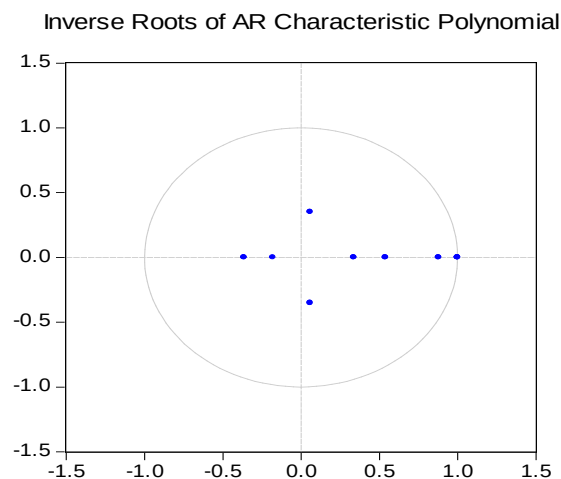


Tests post estimations (EViews)

Corrélation : View/Covariance Analysis... → cocher « Correlation » et « Probability ».

Covariance Analysis: Ordinary Date: 01/13/18 Time: 21:52 Sample: 2002M01 2016M12 Included observations: 180			
Correlation Probability	LIPC	LTXCHANGE	LMM
LIPC	1.000000 -----		
LTXCHANGE	0.987686 0.0000	1.000000 -----	
LMM	0.988314 0.0000	0.971490 0.0000	1.000000 -----

Stabilité : dans l'output de l'estimation, suivre View/Residuals Tests/Lag Structure/AR Roots Graph.



Causalité : dans l'output de l'estimation, suivre View/Residuals Tests/Lag Structure/Granger Causality.Block Exogeneity Tests...

VEC Granger Causality/Block Exogeneity Wald Tests Date: 01/10/18 Time: 19:42 Sample: 2002M01 2016M12 Included observations: 177			
Dependent variable: D(LIPC)			
Excluded	Chi-sq	df	Prob.
D(LTXCHANGE)	2.181726	2	0.3359
D(LMM)	2.090281	2	0.3516
All	4.606577	4	0.3301
Dependent variable: D(LTXCHANGE)			
Excluded	Chi-sq	df	Prob.



D(LIPC)	0.936408	2	0.6261
D(LMM)	2.517914	2	0.2840
All	3.483571	4	0.4804
Dependent variable: D(LMM)			
Excluded	Chi-sq	df	Prob.
D(LIPC)	2.843449	2	0.2413
D(LTXCHANGE)	5.304340	2	0.0705
All	13.52958	4	0.0090

Autocorrélation : dans l'output de l'estimation, suivre View/Residuals Tests/Portementau Autocorrelation Test...

VEC Residual Portmanteau Tests for Autocorrelations

Null Hypothesis: no residual autocorrelations up to lag h					
Date: 01/10/18 Time: 19:44					
Sample: 2002M01 2016M12					
Included observations: 177					
Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	df
1	0.299824	NA*	0.301527	NA*	NA*
2	0.989385	NA*	0.998970	NA*	NA*
3	7.184941	0.6179	7.301345	0.6058	9
4	28.46818	0.0553	29.07668	0.0474	18
5	37.41970	0.0875	38.28842	0.0734	27
6	46.08582	0.1210	47.25861	0.0992	36
7	56.54022	0.1161	58.14349	0.0904	45
8	66.63189	0.1161	68.71288	0.0858	54
9	78.84927	0.0859	81.58475	0.0577	63
10	83.28244	0.1710	86.28338	0.1201	72
11	91.84352	0.1926	95.41176	0.1307	81
12	106.2423	0.1163	110.8577	0.0671	90
*The test is valid only for lags larger than the VAR lag order.					
df is degrees of freedom for (approximate) chi-square distribution					

Autocorrélation : dans l'output de l'estimation, suivre View/Residuals Tests/ Autocorrelation LM Test...

VEC Residual Serial Correlation LM Tests		
Null Hypothesis: no serial correlation at lag order h		
Date: 01/10/18 Time: 19:45		
Sample: 2002M01 2016M12		
Included observations: 177		
Lags	LM-Stat	Prob
1	9.341854	0.4063
2	7.774762	0.5570
3	7.493943	0.5858
4	21.76719	0.0096
5	10.95408	0.2789
6	11.41286	0.2485
7	12.83332	0.1703
8	12.27141	0.1984



9	14.79032	0.0969
10	6.136025	0.7262
11	9.837173	0.3638
12	16.00657	0.0667
Probs from chi-square with 9 df.		

Normalité : dans l'output de l'estimation, suivre View/Residuals Tests/Normality Test...

VEC Residual Normality Tests Orthogonalization: Cholesky (Lutkepohl) Null Hypothesis: residuals are multivariate normal Date: 01/10/18 Time: 19:45 Sample: 2002M01 2016M12 Included observations: 177				
Component	Skewness	Chi-sq	df	Prob.
1	-0.148460	0.650188	1	0.4200
2	0.305193	2.747706	1	0.0974
3	0.126931	0.475288	1	0.4906
Joint		3.873183	3	0.2755
Component	Kurtosis	Chi-sq	df	Prob.
1	14.43192	963.8302	1	0.0000
2	7.879242	175.5767	1	0.0000
3	2.803866	0.283706	1	0.5943
Joint		1139.691	3	0.0000
Component	Jarque-Bera	df	Prob.	
1	964.4803	2	0.0000	
2	178.3244	2	0.0000	
3	0.758995	2	0.6842	
Joint	1143.564	6	0.0000	

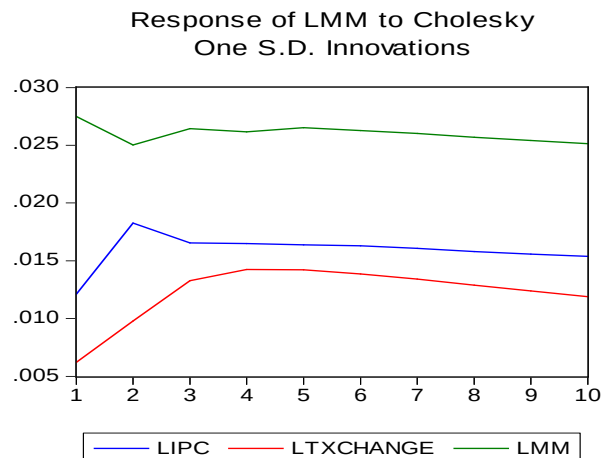
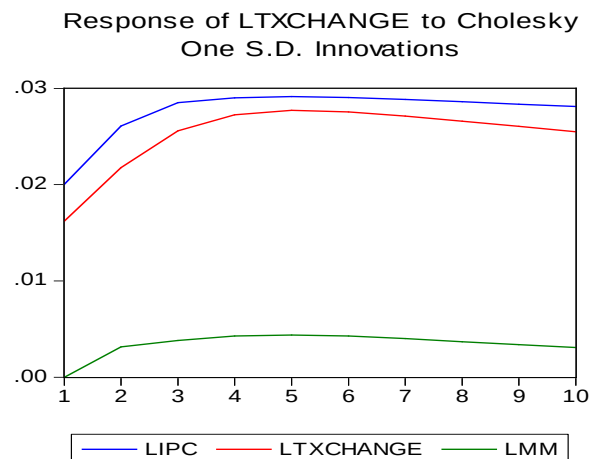
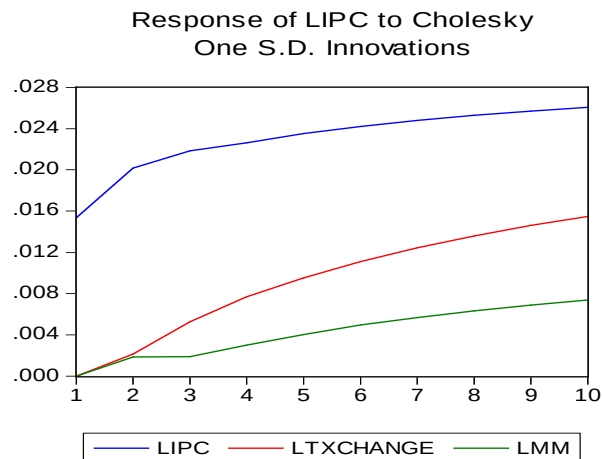
Hétéroscédasticité : dans l'output de l'estimation, suivre View/Residuals Tests/White Heteroskedasticity Test.

VEC Residual Heteroskedasticity Tests: No Cross Terms (only levels and squares)					
Date: 01/10/18 Time: 19:47					
Sample: 2002M01 2016M12					
Included observations: 177					
Joint test:					
Chi-sq	df	Prob.			
233.3396	96	0.0000			
Individual components:					
Dependent	R-squared	F(16,160)	Prob.	Chi-sq(16)	Prob.



res1*res1	0.350434	5.394899	0.0000	62.02685	0.0000
res2*res2	0.317541	4.652890	0.0000	56.20472	0.0000
res3*res3	0.111743	1.257998	0.2306	19.77844	0.2304
res2*res1	0.316244	4.625111	0.0000	55.97528	0.0000
res3*res1	0.222434	2.860647	0.0004	39.37084	0.0010
res3*res2	0.215980	2.754785	0.0006	38.22855	0.0014

Réponses impulsionnelles : dans l'output de l'estimation, cliquer sur l'onglet « impulse », ensuite cocher sur « Combined Graphs ».



Décomposition de la variance : dans l'output de l'estimation, cliquer sur l'onglet « impulse », ensuite cocher sur « Table ».

Variance Decomposition of LIPC:				
Period	S.E.	LIPC	LTXCHANGE	LMM
1	0.015359	100.0000	0.000000	0.000000
2	0.025502	98.75837	0.708774	0.532854
3	0.034042	96.61160	2.779516	0.608885
4	0.041699	93.81653	5.252763	0.930712
5	0.048982	91.04126	7.604764	1.353977
6	0.055981	88.42579	9.756789	1.817424
7	0.062739	86.02633	11.70611	2.267557
8	0.069286	83.84835	13.45762	2.694029
9	0.075645	81.88618	15.02322	3.090594
10	0.081830	80.12282	16.42136	3.455817

Variance Decomposition of LTXCHANGE:				
Period	S.E.	LIPC	LTXCHANGE	LMM
1	0.025791	60.46842	39.53158	0.000000
2	0.042784	59.18205	40.27325	0.544703
3	0.057584	57.25609	42.00064	0.743263
4	0.070151	55.72174	43.40288	0.875389
5	0.080987	54.77801	44.27212	0.949871
6	0.090448	54.24338	44.77254	0.984078
7	0.098820	53.96855	45.04324	0.988207
8	0.106325	53.85807	45.16803	0.973900
9	0.113134	53.85453	45.19657	0.948908
10	0.119377	53.92032	45.16143	0.918248

Variance Decomposition of LMM:				
Period	S.E.	LIPC	LTXCHANGE	LMM
1	0.030671	15.57884	4.102380	80.31878
2	0.044682	24.04763	6.724185	69.22818
3	0.056077	23.97883	9.865402	66.15577
4	0.065598	23.84394	11.92720	64.22887
5	0.074006	23.63261	13.06056	63.30683
6	0.081395	23.55035	13.69612	62.75353
7	0.087972	23.49090	14.04601	62.46309
8	0.093893	23.45858	14.21676	62.32465
9	0.099289	23.44295	14.27058	62.28647
10	0.104256	23.44025	14.24922	62.31053

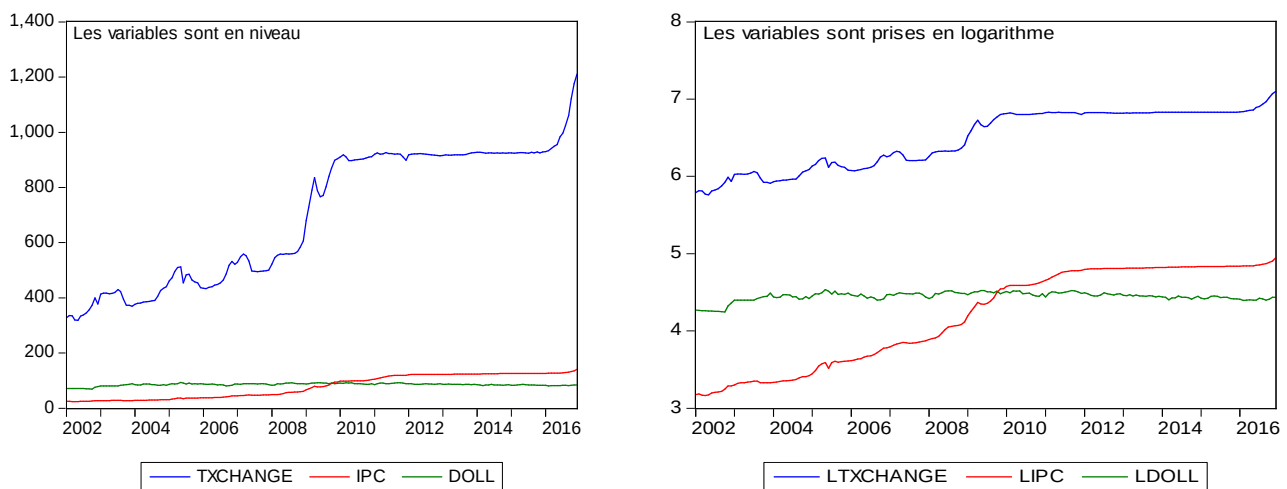
Cholesky Ordering: LIPC LTXCHANGE LMM				
---------------------------------------	--	--	--	--



Cas pratique 3.3 : Vérification de la cointégration et estimation d'un vecteur à correction d'erreur/VECM (taux de change, dollarisation et inflation).

Ici, la démarche est similaire au cas précédent (nous prenons la « dollarisation » à la place de la « masse monétaire »), seulement le souci est d'afficher les équations (de cointégration) d'un VECM pour le cas de plus d'un vecteur de cointégration, soit 2 vecteurs cointégrant le cas d'espèce. Pour ne pas revenir sur les mêmes choses, nous n'allons présenter que les outputs sans commentaires, ni commandes ou chemins de commandes EViews/Stata.

Graphique



Test de stationnarité

Variables	ADF	McKinon à 5%	Stationnarité
Ldollar	-3.299379	-2.877544	I(0)
Lipc	-8.556092	-2.877636	I(1)
Ltxchange	-9.409031	-2.877636	I(1)

Null Hypothesis: LDOLL has a unit root				
Exogenous: Constant				
Lag Length: 0 (Automatic based on SIC, MAXLAG=13)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-3.299379	0.0163
Test critical values:	1% level		-3.466994	
	5% level		-2.877544	
	10% level		-2.575381	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(LDOLL)				
Method: Least Squares				
Date: 01/10/18 Time: 17:20				
Sample (adjusted): 2002M02 2016M12				
Included observations: 179 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.



LDOLL(-1)	-0.078200	0.023702	-3.299379	0.0012
C	0.348674	0.105404	3.307966	0.0011

Décalage optimal

VAR Lag Order Selection Criteria						
Endogenous variables: LDOLL LIPC LTXCHANGE						
Exogenous variables: C						
Date: 01/10/18 Time: 12:20						
Sample: 2002M01 2016M12						
Included observations: 168						
Lag	LogL	LR	FPE	AIC	SC	HQ
0	415.6939	NA	1.48e-06	-4.913022	-4.857237	-4.890382
1	1363.796	1851.057	2.06e-11	-16.09281	-15.86967	-16.00225
2	1387.337	45.12026	1.73e-11*	-16.26592*	-15.87542*	-16.10743*
3	1392.219	9.183654	1.82e-11	-16.21690	-15.65905	-15.99050
4	1395.936	6.857180	1.94e-11	-16.15400	-15.42879	-15.85967
5	1402.772	12.37132	1.99e-11	-16.12824	-15.23568	-15.76600
6	1407.785	8.891605	2.09e-11	-16.08078	-15.02086	-15.65061
7	1415.288	13.04017	2.13e-11	-16.06295	-14.83568	-15.56486
8	1423.041	13.19946	2.17e-11	-16.04811	-14.65348	-15.48210
9	1435.616	20.95812	2.08e-11	-16.09067	-14.52869	-15.45674
10	1441.894	10.23818	2.16e-11	-16.05826	-14.32892	-15.35641
11	1454.142	19.53917*	2.09e-11	-16.09693	-14.20024	-15.32716
12	1463.628	14.79378	2.08e-11	-16.10271	-14.03867	-15.26502
* indicates lag order selected by the criterion						
LR: sequential modified LR test statistic (each test at 5% level)						
FPE: Final prediction error						
AIC: Akaike information criterion						
SC: Schwarz information criterion						
HQ: Hannan-Quinn information criterion						

Test de cointégration de Johansen

Date: 01/10/18 Time: 13:53					
Sample: 2002M01 2016M12					
Included observations: 177					
Series: LDOLL LIPC LTXCHANGE					
Lags interval: 1 to 2					
Selected (0.05 level*) Number of Cointegrating Relations by Model					
Data Trend:	None	None	Linear	Linear	Quadratic
Test Type	No Intercept	Intercept	Intercept	Intercept	Intercept
	No Trend	No Trend	No Trend	Trend	Trend
Trace	2	3	2	1	1
Max-Eig	2	3	2	0	1
*Critical values based on MacKinnon-Haug-Michelis (1999)					
Information Criteria by Rank and Model					
Data Trend:	None	None	Linear	Linear	Quadratic
Rank or	No Intercept	Intercept	Intercept	Intercept	Intercept
No. of CEs	No Trend	No Trend	No Trend	Trend	Trend



Log Likelihood by Rank (rows) and Model (columns)					
0	1412.636	1412.636	1421.511	1421.511	1423.913
1	1422.983	1428.610	1433.562	1433.765	1436.137
2	1432.904	1438.735	1442.411	1442.774	1443.608
3	1432.930	1443.449	1443.449	1443.958	1443.958
Akaike Information Criteria by Rank (rows) and Model (columns)					
0	-15.75860	-15.75860	-15.82498	-15.82498	-15.81822
1	-15.80772	-15.86000	-15.89336	-15.88435	-15.88856
2	-15.85202	-15.89531	-15.92555*	-15.90705	-15.90517
3	-15.78452	-15.86948	-15.86948	-15.84134	-15.84134
Schwarz Criteria by Rank (rows) and Model (columns)					
0	-15.43560	-15.43560	-15.44815*	-15.44815*	-15.38756
1	-15.37705	-15.41140	-15.40886	-15.38190	-15.35023
2	-15.31369	-15.32109	-15.33339	-15.27900	-15.25917
3	-15.13852	-15.16965	-15.16965	-15.08768	-15.08768

Date: 01/10/18 Time: 13:51

Sample (adjusted): 2002M04 2016M12

Included observations: 177 after adjustments

Trend assumption: Linear deterministic trend

Series: LDOLL LIPC LTXCHANGE

Lags interval (in first differences): 1 to 2

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.127310	43.87537	29.79707	0.0007
At most 1 *	0.095152	19.77248	15.49471	0.0106
At most 2	0.011652	2.074511	3.841466	0.1498

Trace test indicates 2 cointegrating eqn(s) at the 0.05 level

* denotes rejection of the hypothesis at the 0.05 level

**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.127310	24.10288	21.13162	0.0185
At most 1 *	0.095152	17.69797	14.26460	0.0138
At most 2	0.011652	2.074511	3.841466	0.1498

Max-eigenvalue test indicates 2 cointegrating eqn(s) at the 0.05 level

* denotes rejection of the hypothesis at the 0.05 level

**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegrating Coefficients (normalized by b*S11*b=l):

LDOLL	LIPC	LTXCHANGE
22.56740	1.170736	-2.854256
-1.557723	-12.12120	19.06851
-4.346497	-3.019694	7.760961

Unrestricted Adjustment Coefficients (alpha):

D(LDOLL)	-0.005388	0.003113	0.000179
D(LIPC)	0.000923	0.001625	-0.001549
D(LTXCHANGE)	-0.001637	-0.002112	-0.002500



1 Cointegrating Equation(s):		Log likelihood	1433.562
Normalized cointegrating coefficients (standard error in parentheses)			
LDOLL	LIPC	LTXCHANGE	
1.000000	0.051877	-0.126477	
	(0.10966)	(0.17875)	
Adjustment coefficients (standard error in parentheses)			
D(LDOLL)	-0.121592		
	(0.03025)		
D(LIPC)	0.020838		
	(0.02687)		
D(LTXCHANGE)	-0.036936		
	(0.04259)		

2 Cointegrating Equation(s):		Log likelihood	1442.411
Normalized cointegrating coefficients (standard error in parentheses)			
LDOLL	LIPC	LTXCHANGE	
1.000000	0.000000	-0.045167	
		(0.02458)	
0.000000	1.000000	-1.567349	
		(0.05365)	
Adjustment coefficients (standard error in parentheses)			
D(LDOLL)	-0.126440	-0.044037	
	(0.02983)	(0.01606)	
D(LIPC)	0.018306	-0.018618	
	(0.02679)	(0.01442)	
D(LTXCHANGE)	-0.033646	0.023685	
	(0.04253)	(0.02289)	

Estimation du VECM

Basics Cointegration VEC Restrictions

VAR Type

☐ Unrestricted VAR

☒ Vector Error Correction

Endogenous Variables

ltxchange ldoll lipc

Estimation Sample

2002m01 2016m12

Lag Intervals for D(Endogenous):

1 2

Exogenous Variables

Do NOT include C or Trend in VEC's

OK Annuler

Basics Cointegration VEC Restrictions

Rank

Number of cointegrating

Deterministic Trend Specification

No trend in data

☐ 1) No intercept or trend in CE or VAR

☐ 2) Intercept (no trend) in CE - no intercept in VAR

Linear trend in data

☒ 3) Intercept (no trend) in CE and VAR

☐ 4) Intercept and trend in CE - no trend in VAR

Quadratic trend in data

☐ 5) Intercept and trend in CE - linear trend in VAR

OK Annuler



Vector Error Correction Estimates Date: 01/10/18 Time: 14:15 Sample (adjusted): 2002M04 2016M12 Included observations: 177 after adjustments Standard errors in () & t-statistics in []			
Cointegrating Eq:	CointEq1	CointEq2	
LTXCHANGE(-1)	1.000000	0.000000	
LDOLL(-1)	0.000000	1.000000	
LIPC(-1)	-0.638020 (0.02106) [-30.2939]	-0.028818 (0.01532) [-1.88157]	
C	-3.810189	-4.327781	
Error Correction:	D(LTXCHANGE)	D(LDOLL)	D(LIPC)
CointEq1	-0.035603 (0.03636) [-0.97926]	0.074732 (0.02550) [2.93053]	0.028355 (0.02290) [1.23821]
CointEq2	-0.033646 (0.04266) [-0.78879]	-0.126440 (0.02992) [-4.22611]	0.018306 (0.02687) [0.68137]
D(LTXCHANGE(-1))	0.327272 (0.12058) [2.71413]	-0.013242 (0.08458) [-0.15657]	0.054200 (0.07595) [0.71364]
D(LTXCHANGE(-2))	0.157910 (0.11887) [1.32839]	-0.060432 (0.08338) [-0.72480]	0.069584 (0.07487) [0.92936]
D(LDOLL(-1))	-0.280170 (0.10499) [-2.66864]	-0.111181 (0.07364) [-1.50984]	-0.100069 (0.06613) [-1.51331]
D(LDOLL(-2))	0.124031 (0.10439) [1.18812]	-0.051549 (0.07322) [-0.70401]	0.051377 (0.06575) [0.78137]
D(LIPC(-1))	0.002431 (0.18929) [0.01284]	0.256313 (0.13277) [1.93047]	0.277712 (0.11923) [2.32925]
D(LIPC(-2))	-0.053505 (0.19037) [-0.28105]	0.106496 (0.13353) [0.79755]	-0.019742 (0.11991) [-0.16464]
C	0.004515 (0.00246) [1.83543]	-0.001871 (0.00173) [-1.08427]	0.006732 (0.00155) [4.34501]
R-squared	0.180060	0.202422	0.197581
Adj. R-squared	0.141015	0.164442	0.159370
Sum sq. resids	0.105732	0.052017	0.041946
S.E. equation	0.025087	0.017596	0.015801
F-statistic	4.611623	5.329722	5.170852
Log likelihood	405.7830	468.5597	487.6044



Akaike AIC	-4.483424	-5.192765	-5.407960
Schwarz SC	-4.321925	-5.031266	-5.246461
Mean dependent	0.007264	0.000985	0.010049
S.D. dependent	0.027068	0.019250	0.017234
Determinant resid covariance (dof adj.)		1.96E-11	
Determinant resid covariance		1.68E-11	
Log likelihood		1442.411	
Akaike information criterion		-15.92555	
Schwarz criterion		-15.33339	

Estimation Proc:

```
=====
EC(C,2) 1 2 LTXCHANGE LDOLL LIPC
```

VAR Model:

```
=====
D(LTXCHANGE) = A(1,1)*(B(1,1)*LTXCHANGE(-1) + B(1,2)*LDOLL(-1) + B(1,3)*LIPC(-1) + B(1,4)) +
A(1,2)*(B(2,1)*LTXCHANGE(-1) + B(2,2)*LDOLL(-1) + B(2,3)*LIPC(-1) + B(2,4)) + C(1,1)*D(LTXCHANGE(-1)) +
C(1,2)*D(LTXCHANGE(-2)) + C(1,3)*D(LDOLL(-1)) + C(1,4)*D(LDOLL(-2)) + C(1,5)*D(LIPC(-1)) +
C(1,6)*D(LIPC(-2)) + C(1,7)
```

```
D(LDOLL) = A(2,1)*(B(1,1)*LTXCHANGE(-1) + B(1,2)*LDOLL(-1) + B(1,3)*LIPC(-1) + B(1,4)) +
A(2,2)*(B(2,1)*LTXCHANGE(-1) + B(2,2)*LDOLL(-1) + B(2,3)*LIPC(-1) + B(2,4)) + C(2,1)*D(LTXCHANGE(-1)) +
C(2,2)*D(LTXCHANGE(-2)) + C(2,3)*D(LDOLL(-1)) + C(2,4)*D(LDOLL(-2)) + C(2,5)*D(LIPC(-1)) +
C(2,6)*D(LIPC(-2)) + C(2,7)
```

```
D(LIPC) = A(3,1)*(B(1,1)*LTXCHANGE(-1) + B(1,2)*LDOLL(-1) + B(1,3)*LIPC(-1) + B(1,4)) +
A(3,2)*(B(2,1)*LTXCHANGE(-1) + B(2,2)*LDOLL(-1) + B(2,3)*LIPC(-1) + B(2,4)) + C(3,1)*D(LTXCHANGE(-1)) +
C(3,2)*D(LTXCHANGE(-2)) + C(3,3)*D(LDOLL(-1)) + C(3,4)*D(LDOLL(-2)) + C(3,5)*D(LIPC(-1)) +
C(3,6)*D(LIPC(-2)) + C(3,7)
```

VAR Model - Substituted Coefficients:

```
=====
D(LTXCHANGE) = - 0.0356031839097*( LTXCHANGE(-1) - 0.638020070146*LIPC(-1) - 3.81018940117 ) -
0.0336461175068*( LDOLL(-1) - 0.0288175158592*LIPC(-1) - 4.32778093141 ) +
0.327272482275*D(LTXCHANGE(-1)) + 0.157909782925*D(LTXCHANGE(-2)) - 0.280169529306*D(LDOLL(-1))
+ 0.124031464763*D(LDOLL(-2)) + 0.00243068796979*D(LIPC(-1)) - 0.0535051929405*D(LIPC(-2)) +
0.00451462703222
```

```
D(LDOLL) = 0.0747317111534*( LTXCHANGE(-1) - 0.638020070146*LIPC(-1) - 3.81018940117 ) -
0.12644033583*( LDOLL(-1) - 0.0288175158592*LIPC(-1) - 4.32778093141 ) -
0.0132418001036*D(LTXCHANGE(-1)) - 0.0604324405551*D(LTXCHANGE(-2)) - 0.111181026926*D(LDOLL(-
1)) - 0.051548783578*D(LDOLL(-2)) + 0.256313236611*D(LIPC(-1)) + 0.106496063015*D(LIPC(-2)) -
0.00187064301056
```

```
D(LIPC) = 0.0283545735992*( LTXCHANGE(-1) - 0.638020070146*LIPC(-1) - 3.81018940117 ) +
0.0183062351675*( LDOLL(-1) - 0.0288175158592*LIPC(-1) - 4.32778093141 ) +
0.054200116847*D(LTXCHANGE(-1)) + 0.0695841459942*D(LTXCHANGE(-2)) - 0.100068984044*D(LDOLL(-
1)) + 0.0513771552532*D(LDOLL(-2)) + 0.277711801463*D(LIPC(-1)) - 0.01974178851*D(LIPC(-2)) +
0.0067315371403
```



CHAP IV : MODELES LINEAIRES A EQUATIONS SIMULTANEEES

I. Estimation des modèles à équations simultanées

Cas pratique 4.1 : Modèle de Consommation

► **Modèle** : considérons le modèle linéaire à équations simultanées (traduisant la consommation) ci-dessous :

$$\begin{cases} CO_t = \beta_0 + \beta_1 R_t + e_t \dots \dots (1) \\ R_t = CO_t + Z_t \dots \dots \dots (2) \end{cases} \dots \dots [4.1]$$

Avec :

- o CO_t : variable endogène (Dépenses en consommation observées au temps t) ;
- o R_t : Variable endogène (Revenu observé au temps t) ;
- o Z_t : variable exogène (dépenses en investissement observées au temps t).

► **Travail demandé** :

- Identifier l'équation structurelle (1) ;
- Estimer les équations structurelles (1) et (2) par les Double Moindres Carrés/DMC et les Moindre Carrés Indirects/MCI, et comparer les résultats ;
- Produire les dépenses en consommation attendues pour les années 2000, 2001 et 2002.

a) Etude de l'identification de « l'équation (1) »

_____ Rappels sur les règles d'identification :

Soient les abréviations suivantes :

$$\left\{ \begin{array}{l} g : \text{nombre de variables endogènes (équations) dans le système/modèle ;} \\ g' : \text{nombre de variables endogènes présentes dans l'équation à identifier ;} \\ k : \text{nombre de variables exogènes dans le système/modèle ;} \\ k' : \text{nombre de variables exogènes figurant dans l'équation à identifier ;} \\ r : \text{nombre de restrictions figurant dans l'équation à identifier (restreindre un paramètre structurel c'est lui fixer une valeur suivant l'écriture du modèle)} \end{array} \right.$$

Les règles de décision sur le statut (l'identifiabilité) d'une équation sont les suivantes :

- o $(g - g') + (k - k') + r = g - 1$: l'équation est juste identifiée ;
- o $(g - g') + (k - k') + r > g - 1$: l'équation est sur-identifiée ;
- o $(g - g') + (k - k') + r < g - 1$: l'équation est non ou sous identifiée.



Identification de l'équation (1) : avec : $g=2$ (CO_t , R_t); $g'=2$ (CO_t , R_t); $k=1$ (Z_t); $k'=0$ et $r=0$, nous affirmons que cette équation est juste identifiée, car :

$$(g - g') + (k - k') + r = g - 1 \Rightarrow (2 - 2) + (1 - 0) + 0 = 2 - 1 \Leftrightarrow 1 = 1$$

b) Estimation des équations structurelles (1) et (2) par les Double Moindres Carrés/DMC et les Moindre Carrés Indirects/MCI, et comparaison des résultats

► **Les MCI :**

Modèle sous forme réduite :

(2) dans (1) $\Rightarrow CO_t = \beta_0 + \beta_1(CO_t + Z_t) + e_t$, ce qui nous permet d'écrire :

$$CO_t = \frac{\beta_0}{(1 - \beta_1)} + \left(\frac{\beta_0}{(1 - \beta_1)} \right) * Z_t + \frac{e_t}{(1 - \beta_1)} \dots \dots \dots [3]$$

(3) dans (2) donne :

$$R_t = \frac{\beta_0}{(1 - \beta_1)} + \left(\frac{1}{(1 - \beta_1)} \right) * Z_t + \frac{e_t}{(1 - \beta_1)} \dots \dots \dots [4]$$

Posons : $\alpha_0 = \frac{\beta_0}{(1 - \beta_1)}$; $\alpha_1 = \frac{1}{(1 - \beta_1)}$ et $v_t = \frac{e_t}{(1 - \beta_1)}$, ce qui implique les équations réduites suivantes :

$$\begin{cases} CO_t = \alpha_0 + (\alpha_1 - 1)Z_t + v_t \dots \dots (1') \\ R_t = \alpha_0 + \alpha_1 Z_t + v_t \dots \dots \dots (2') \end{cases} \dots \dots \dots [4.2]$$

Estimation de « l'équation (2') » par les MCO :

Sur Eviews :	Sur Stata, faire:
<pre>{ create a 1993 2002 data COt Zt Rt ls Rt c Zt</pre>	<pre>{ tsset annee reg Rt Zt</pre>

Source	SS	df	MS	Number of obs = 7		
Model	1452.86429	1	1452.86429	F(1, 5) =	142.06	
Residual	51.1357143	5	10.2271429	Prob > F =	0.0001	
Total	1504	6	250.666667	R-squared =	0.9660	
				Adj R-squared =	0.9592	
				Root MSE =	3.198	
Rt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
Zt	3.221429	.2702795	11.92	0.000	2.526653	3.916204
_cons	46.02143	7.396973	6.22	0.002	27.0069	65.03595

De cette estimation, nous obtenons les paramètres réduits estimés : $\hat{\alpha}_0 = 46.02$ et $\hat{\alpha}_1 = 3.22$, grâce auxquels nous allons déduire les paramètres structurels suivant les relations algébriques : $\alpha_0 = \frac{\beta_0}{(1 - \beta_1)}$; $\alpha_1 = \frac{1}{(1 - \beta_1)}$. Il est pour nous question de calculer β_0 et β_1 . Pour ce faire, écrivons :



$$\text{D'un côté : } 3.22 = \frac{1}{(1-\beta_1)} \Rightarrow (1-\beta_1) = \frac{1}{3.22} \Leftrightarrow (1-\beta_1) = 0.31055 \Rightarrow \beta_1 = 0.69$$

$$\text{De l'autre : } 46.02 = \frac{\beta_0}{(1-\beta_1)} \Rightarrow \beta_0 = 46.02 * 0.31055 = 14.29$$

► **Les DMC :**

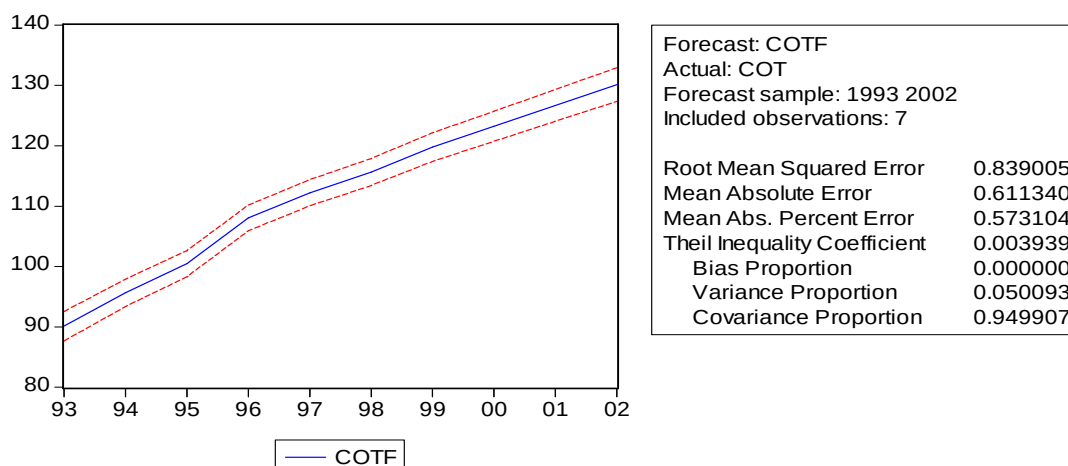
Sur Eviews, faire : **TSLS COT c Rt @ c Zt**

Dependent Variable: COT				
Method: Two-Stage Least Squares				
Date: 12/14/13 Time: 01:51				
Sample (adjusted): 1993 1999				
Included observations: 7 after adjustments				
Instrument list: C ZT				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	14.28603	3.484182	4.100254	0.0094
RT	0.689579	0.026045	26.47694	0.0000
R-squared	0.993359	Mean dependent var		106.0000
Adjusted R-squared	0.992031	S.D. dependent var		11.12055
S.E. of regression	0.992724	Sum squared resid		4.927508
F-statistic	701.0281	Durbin-Watson stat		2.362376
Prob(F-statistic)	0.000001			

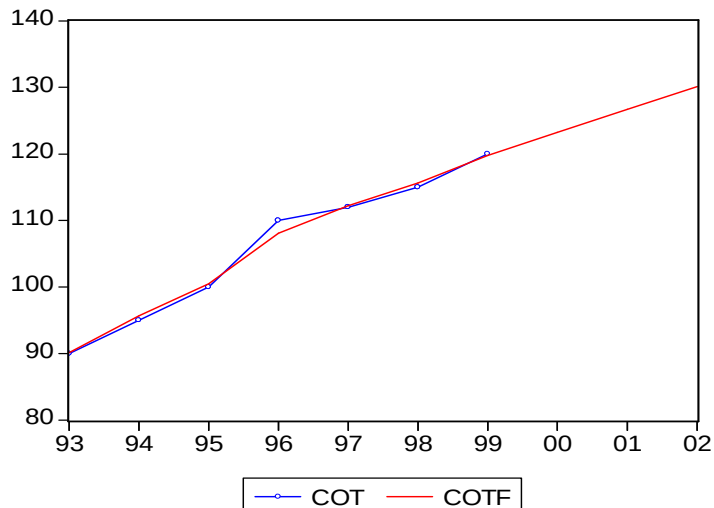
► **Commentaires :** pour un modèle juste identifié, les estimateurs des MCI sont similaires à ceux obtenus par les DMC (en pratique, les DMC sont préférables aux MCI du fait de la simplicité de sa mise en œuvre/ses procédures d'estimation). **NB :** $R^2=99.34\%$ (le revenu/RT influe significativement sur la consommation/COT : un revenu supplémentaire de 10 unités monétaires occasionne une propension marginale à consommer de 6.90 points).

c) **Dépenses en consommation attendues pour les années 2000, 2001 et 2002**

Sur Eviews, pour prévoir les dépenses en consommation, dans l'output, cliquer sur l'onglet **Forecast** → ok (avec COTf = COT prévue ou ajustée) :



Commentaire : au regard de l'indice d'inégalité de Theil (proche de zéro) et du MAPE qui est à 57.3% (Mean Absolute Percent Error), nous concluons en faveur d'une prévision de bonne qualité. Le graphique affichant les valeurs observées et prévues est illustratif (sur Eviews, taper : **plot COT COTf** pour le graphique et **show COT COTf** pour les valeurs sous forme tabulaire) :



obs	COT	COTF
1993	90.00000	90.13969
1994	95.00000	95.65632
1995	100.0000	100.4834
1996	110.0000	108.0687
1997	112.0000	112.2062
1998	115.0000	115.6541
1999	120.0000	119.7916
2000	NA	123.2395
2001	NA	126.6874
2002	NA	130.1353

Cas pratique 4.2 : Modèle macroéconomique simple

► **Modèle :** Soit le modèle linéaire à équations simultanées suivant (les données sont centrées, ce qui justifie l'absence de termes constants) :

$$\begin{cases} Y_{1t} = a(Y_{2t} + X_{1t}) + e_{1t} \dots \dots (5) \\ \dots \dots \dots [4.3] \\ Y_{2t} = bY_{1t} + cY_{t-1} + e_{2t} \dots \dots (6) \end{cases}$$

Avec :

- o Y_{1t} : variable endogène (Production nationale au temps t) ;
- o Y_{2t} : Variable endogène (Consommation de ménages au temps t) ;
- o X_{1t} : variable exogène (demande finale autre que Y_{2t} au temps t) ;
- o Y_{1t-1} : variable prédéterminé ou exogène (production de l'an passé).

► **Travail demandé :**

- Identifier le modèle ;
- Estimer les équations structurelles (5) et (6) par les Moindres Carrés Ordinaires/MCO ;
- Estimer les équations (5) et (6) par les Double Moindres Carrés/DMC et les Moindre Carrés Indirects/MCI ;
- Comparer les résultats.



a) Identification du modèle**NB :** $g=2$ (Y_{1t} et Y_{2t}) et $k=2$ (X_{1t} et Y_{1t-1})

Equations	Identification	Conclusion	Méthode
Eq.(5)	$g'=2$ (Y_{1t} et Y_{2t}) ; $k'=1$ (X_{1t}) et $r=1$ $(Y_{1t}$ et Y_{2t} ont les mêmes coefficients). $(g-g')+(k-k')+r ? (g-1)$ $(2-2)+(2-1)+1 > 2-1$	L'équation (5) est sur-identifiée	Doubles Moindres Carrés (DMC)
Eq.(6)	$g'=2$ (Y_{1t} et Y_{2t}) ; $k'=1$ (Y_{1t-1}) et $r=0$. $(g-g')+(k-k')+r ? (g-1)$ $(2-2)+(2-1)+0 = 2-1$	L'équation (6) est juste identifiée	Doubles Moindres Carrés/DMC ou Moindres Carrés Indirects/MCI

b) Estimation des équations structurelles (5) et (6) par les Moindres Carrés Ordinaires/MCO

Sur Stata, faire :

```

{
tsset time
gen YS=Y2t+X1t
reg Y1t YS,noconstant
reg Y2t Y1t Y1LAG,nocons

```

Source	SS	df	MS	Number of obs = 9		
Model	2274.36667	1	2274.36667	F(1, 8)	=	34.75
Residual	523.633333	8	65.4541667	Prob > F	=	0.0004
Total	2798	9	310.888889	R-squared	=	0.8129
				Adj R-squared	=	0.7895
				Root MSE	=	8.0904
Y1t	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
YS	.5166667	.0876494	5.89	0.000	.3145468	.7187865

Source	SS	df	MS	Number of obs = 9		
Model	777.036285	2	388.518142	F(2, 7)	=	3.56
Residual	762.963715	7	108.994816	Prob > F	=	0.0856
Total	1540	9	171.111111	R-squared	=	0.5046
				Adj R-squared	=	0.3630
				Root MSE	=	10.44
Y2t	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
Y1t	.5187397	.2010013	2.58	0.036	.0434471	.9940323
Y1LAG	.0402599	.2189994	0.18	0.859	-.4775915	.5581112

NB : Dans le cas d'un système multi-équationnel, les estimateurs de MCO sont biaisés.

c) Estimation des équations (5) et (6) par les Doubles Moindres Carrés/DMC et les Moindres Carrés Indirects/MCI

► **Estimation par les MCI :**

Sous forme réduite, les expressions (5) et (6) deviennent :

$$\begin{cases} Y_{1t} = \left(\frac{a}{1-ab}\right) X_{1t} + \left(\frac{ac}{1-ab}\right) Y_{1t-1} + \left(\frac{1}{1-ab}\right) (e_{1t} + e_{2t}) \dots \dots (5') \\ Y_{2t} = \left(\frac{ab}{1-ab}\right) X_{1t} + \left(\frac{c}{1-ab}\right) Y_{1t-1} + \left(\frac{1}{1-ab}\right) (be_{1t} + e_{2t}) \dots \dots (6') \end{cases} \dots \dots [4.4]$$

Posons : $\pi_0 = \left(\frac{a}{1-ab}\right)$; $\pi_1 = \left(\frac{ac}{1-ab}\right)$; $\pi_2 = \left(\frac{ab}{1-ab}\right)$; $\pi_3 = \left(\frac{c}{1-ab}\right)$;

$$u_{1t} = \left(\frac{1}{1-ab}\right) (e_{1t} + e_{2t}) \text{ et } u_{2t} = \left(\frac{1}{1-ab}\right) (be_{1t} + e_{2t})$$

Ce qui nous amène à écrire :

$$\begin{cases} Y_{1t} = \pi_0 X_{1t} + \pi_1 Y_{1t-1} + u_{1t} \dots \dots (5'') \\ Y_{2t} = \pi_2 X_{1t} + \pi_3 Y_{1t-1} + u_{2t} \dots \dots (6'') \end{cases} \dots \dots [4.4']$$

Les résultats d'estimation des expressions (5'') et (6'') se présentent comme suit :

Commande: reg Y1t X1t Y1LAG, nocons

Source	SS	df	MS	Number of obs = 9		
Model	2195.656	2	1097.828	F(2, 7)	=	12.76
Residual	602.344005	7	86.0491435	Prob > F	=	0.0046
Total	2798	9	310.888889	R-squared	=	0.7847
				Adj R-squared	=	0.7232
				Root MSE	=	9.2763
Y1t	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
X1t	.7174487	.1453871	4.93	0.002	.3736628	1.061235
Y1LAG	.1909748	.1910953	1.00	0.351	-.2608938	.6428434

Commande: reg Y2t X1t Y1LAG, nocons

Source	SS	df	MS	Number of obs = 9		
Model	565.159746	2	282.579873	F(2, 7)	=	2.03
Residual	974.840254	7	139.262893	Prob > F	=	0.2018
Total	1540	9	171.111111	R-squared	=	0.3670
				Adj R-squared	=	0.1861
				Root MSE	=	11.801
Y2t	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
X1t	.3553575	.1849568	1.92	0.096	-.0819959	.792711
Y1LAG	.1396827	.2431053	0.57	0.584	-.4351701	.7145355



► **Estimation par les DMC :**

Sur Eviews, taper :

TSLS Y1t YS @ X1t Y1LAG : estimer l'équation (5) ;

TSLS Y2t Y1t Y1LAG @ X1t Y1LAG : estimer l'équation (6) ;

Dependent Variable: Y1T Method: Two-Stage Least Squares Date: 12/14/13 Time: 05:26 Sample: 1995 2003 Included observations: 9 Instrument list: X1T Y1LAG				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
YS	0.535520	0.093407	5.733162	0.0004
R-squared	0.811772	Mean dependent var		0.000000
Adjusted R-squared	0.811772	S.D. dependent var		18.70160
S.E. of regression	8.113737	Sum squared resid		526.6618
Durbin-Watson stat	1.894670			

Dependent Variable: Y2T Method: Two-Stage Least Squares Date: 12/14/13 Time: 05:30 Sample: 1995 2003 Included observations: 9 Instrument list: X1T Y1LAG				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
Y1T	0.495344	0.228271	2.169985	0.0666
Y1LAG	0.045084	0.220336	0.204614	0.8437
R-squared	0.503610	Mean dependent var		0.000000
Adjusted R-squared	0.432697	S.D. dependent var		13.87444
S.E. of regression	10.45016	Sum squared resid		764.4403
Durbin-Watson stat	3.011822			

d) **Comparaison des résultats**

Paramètres	METHODES		
	MCO	MCI	DMC
$\hat{a}$	0.517/ S	$\hat{\pi}_0 = 0.717/S ; \hat{\pi}_1 = 0.191/NS ;$ $\hat{\pi}_2 = 0.355/NS ; \hat{\pi}_3 = 0.140/NS$	0.536/ S
$\hat{b}$	0.519/ S	0.495	0.495/ NS
$\hat{c}$	0.040/ NS	0.045	0.045/ NS

Note : Bien que significatifs, les estimateurs des MCO ne sont pas BLUE pour un système d'équations simultanées (il y a biais : surestimation ou sous-estimation des paramètres). Pour l'expression (6) juste identifiée, le constat est que les résultats sont les mêmes (estimateurs identiques comme nous l'avons souligné au cas précédent) ; toutefois, en pratique, les DMC sont préférables au MCI.



Cas pratique 4.3 : Modèle macroéconomique de KLEIN (1950)

- **Modèle** : Soit le modèle macroéconomique de KLEIN (il s'agit d'un système linéaire à équations simultanées avec 6 équations et 11 variables économiques, en plus du vecteur « unité ») suivant :

$$\begin{cases} CO_t = \pi_0 + \pi_1 P_t + \pi_2 P_{t-1} + \pi_3 (W_{it} + W_{at}) + e_{1t} \dots \dots (7) \\ I_t = \gamma_0 + \gamma_1 P_t + \gamma_2 P_{t-1} + \gamma_3 K_{t-1} + e_{2t} \dots \dots \dots (8) \\ W_{it} = \alpha_0 + \alpha_1 X_t + \alpha_2 X_{t-1} + \alpha_3 t + e_{3t} \dots \dots \dots (9) \dots \dots \dots [4.5] \\ X_t = CO_t + I_t + G_t \dots \dots \dots (10) \\ P_t = X_t - W_{it} - T_t \dots \dots \dots (11) \\ K_t = I_t + K_{t-1} \dots \dots \dots (12) \end{cases}$$

Avec :

- o CO_t : Consommation au temps t ;
- o P_t : Profits au temps t ;
- o W_{it} : Salaires (wage) versés par le secteur « industrie » au temps t ;
- o W_{at} : Salaires versés par le secteur « administration » au temps t ;
- o I_t : Investissement au temps t ;
- o X_t : Production industrielle au temps t ;
- o K_t : Stock de capital à la fin de l'année t ;
- o T_t : Impôt sur les bénéfices (profits) au temps t ;
- o G_t : Dépenses publiques au temps t ;
- o t : le temps (la tendance/trend)
- o $K_{t-1} = KLAG$.

► **Travail demandé :**

- Faire une représentation graphique des trajectoires des variables sous-étude ;
- Repérer les variables endogènes à distinguer de variables exogènes et étudier l'identification de ce modèle ;
- Estimer les équations structurelles (7), (8) et (9) par les Moindres Carrés Ordinaires/MCO ;
- Recourir à la méthode adéquate pour estimer le modèle.
- Résoudre le modèle et effectuer une prévision.

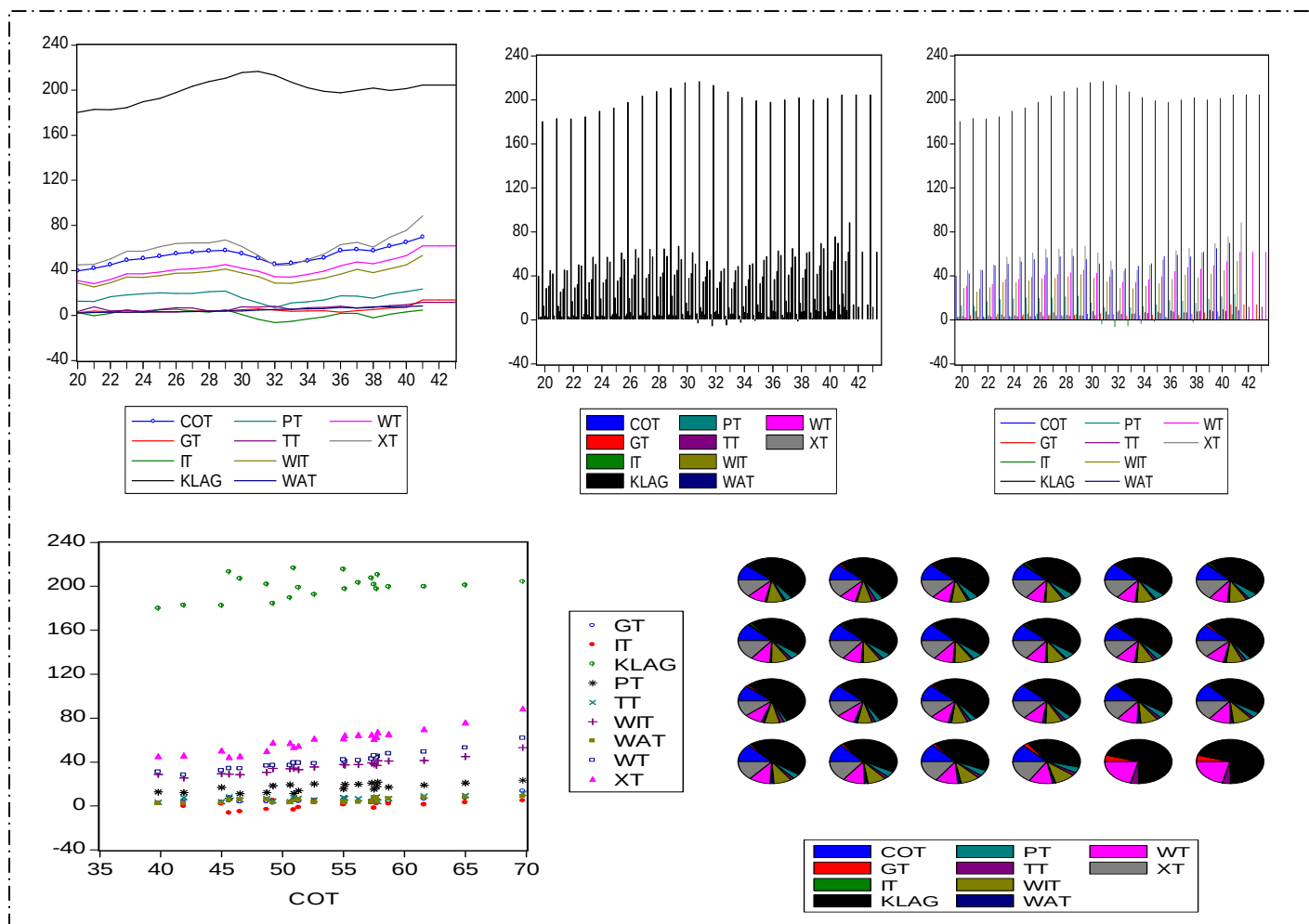
a) représentation graphique des trajectoires des variables sous-étude

Sur Eviews, dans l'output des données, suivre : View/Graph → Choisir ensuite (pour les variables groupées ou prises ensemble) :

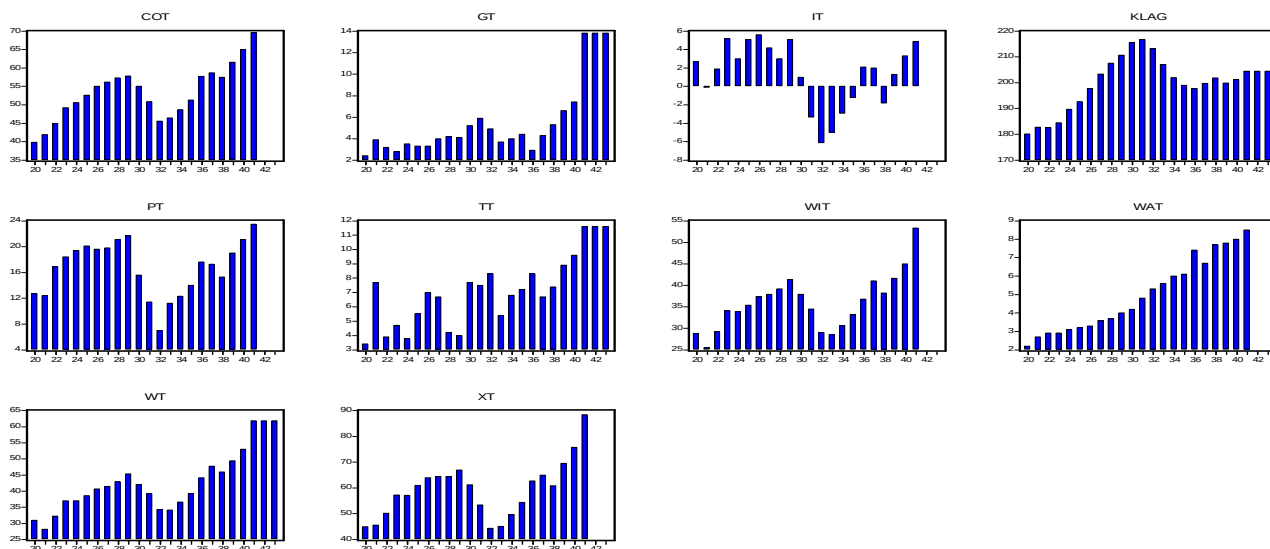
- o line : pour les lignes ;
- o Are : pour les surfaces (aires) ;
- o Bar : pour les barres (bâtons) ;
- o Spike : pour les pointes (tige) ;
- o Scatter/Simple Scatter : pour le nuage de points ;



o Pie : pour les gâteaux.



Sur EViews, dans l'output des données, suivre : View/Multiple Graphs→ Choisir ensuite (pour les variables prises séparément) : line, Are, **Bar** ou Spike.



b) Variables endogènes vs variables exogènes (distinction) et Identification du modèle

► Distinction des variables endogènes et exogènes – notons ce qui suit :

- o Les variables endogènes sont (6) : COt, It, WIt, Xt, Pt et Kt (**g = 6**) ;
- o Les variables exogènes sont (8) : Gt, WAt, Tt, t, Xt-1, Pt-1, Kt-1 et le vecteur de constantes ou vecteur unité (**k = 8**).

► Etude de conditions d'identification du modèle : Seules les équations de comportement ou techniques (7), (8) et (9) font l'objet de l'étude des conditions d'identification. Les expressions (10), (11) et (12), traduisant des identités comptables (relations d'équilibre), ne peuvent en aucun cas faire l'objet d'une quelconque identification (*aucun paramètre n'est à estimer*).

Modèle	Identification	Conclusion	Méthode
Equat (7)	$g'=3$ (COt, Pt et Wit) ; $k'=3$ (Pt-1, WAt, constante) et $r=1$ (Wit et WAt ont des coefficients identiques). $(g-g')+(k-k')+r ? (g-1)$ $(6-3)+(8-3)+1 > 6-1$	L'équation (7) est sur-identifiée	Doubles Moindres Carrés (DMC)
Equat (8)	$g'=2$ (It et Pt) ; $k'=3$ (Pt-1, Kt-1 et constante) et $r=0$. $(g-g')+(k-k')+r ? (g-1)$ $(6-2)+(8-3)+0 > 6-1$	L'équation (8) est sur-identifiée	Doubles Moindres Carrés (DMC)
Equat (9)	$g'=2$ (Wit et Xt) ; $k'=3$ (Xt-1, t et constante) et $r=0$. $(g-g')+(k-k')+r ? (g-1)$ $(6-2)+(8-3)+0 > 6-1$	L'équation (9) est sur-identifiée	Doubles Moindres Carrés (DMC)

c) Estimation des équations structurelles (7), (8) et (9) par les Moindres Carrés Ordinaires/MCO

_____ Sur Stata, faire :

```

{
tsset annee
reg COt Pt L.Pt Wt
reg It Pt L.Pt Klag
reg WIt Xt L.Xt Klag
}

```

_____ Sur Eviews, Faire :

```

create a 1920 1943
data COt Gt It Klag Pt Tt Wit WAt Wt Xt
genr t=@trend(1)
ls COt c Pt Pt(-1) Wt
Equation eq1.ls COt c Pt Pt(-1) Wt
Equation eq2.ls It c Pt Pt(-1) Klag
Equation eq3.ls WIt c Xt Xt(-1) t

```

Les résultats se présentent comme suit:

```

COt = 16.23660027 + 0.1929343813*Pt + 0.08988489781*Pt(-1) + 0.7962187497*Wt
It = 10.12578854 + 0.4796356446*Pt + 0.3330387135*Pt(-1) - 0.1117946837*KLAg
WIt = 10.61420996 + 0.4394769672*Xt + 0.1460899468*Xt(-1) + 0.1302452303*t

```

Note : on peut écrire : Equation eq1.ls COt c Pt Pt(-1) (**Wit+Wat**)



Commande: reg C0t Pt L.Pt Wt

Source	SS	df	MS	Number of obs = 21		
Model	923.549941	3	307.84998	F(3, 17)	=	292.71
Residual	17.8794481	17	1.05173224	Prob > F	=	0.0000
Total	941.429389	20	47.0714695	R-squared	=	0.9810
				Adj R-squared	=	0.9777
				Root MSE	=	1.0255

C0t	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
Pt						
--.	.1929342	.0912102	2.12	0.049	.0004976	.3853709
L1.	.0898849	.0906479	0.99	0.335	-.1013655	.2811354
Wt	.7962187	.0399439	19.93	0.000	.7119444	.880493
_cons	16.2366	1.302698	12.46	0.000	13.48815	18.98506

Commande: reg It Pt L.Pt Klag

Source	SS	df	MS	Number of obs = 21		
Model	235.00396	3	78.3346533	F(3, 17)	=	76.88
Residual	17.3226985	17	1.01898226	Prob > F	=	0.0000
Total	252.326659	20	12.6163329	R-squared	=	0.9313
				Adj R-squared	=	0.9192
				Root MSE	=	1.0094

It	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
Pt						
--.	.4796356	.0971146	4.94	0.000	.2747418	.6845294
L1.	.3330387	.1008592	3.30	0.004	.1202444	.545833
Klag	-.1117947	.0267276	-4.18	0.001	-.1681849	-.0554045
_cons	10.12579	5.465546	1.85	0.081	-1.405502	21.65709

Commande: reg WIt Xt L.Xt t

Source	SS	df	MS	Number of obs = 21		
Model	784.904754	3	261.634918	F(3, 17)	=	444.57
Residual	10.0047374	17	.588513967	Prob > F	=	0.0000
Total	794.909491	20	39.7454746	R-squared	=	0.9874
				Adj R-squared	=	0.9852
				Root MSE	=	.76715

WIt	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
Xt						
--.	.4394769	.0324076	13.56	0.000	.371103	.5078509
L1.	.14609	.0374231	3.90	0.001	.0671341	.2250458
t	.1302452	.0319103	4.08	0.001	.0629203	.19757
_cons	.0643462	1.151797	0.06	0.956	-2.365732	2.494425

d) Estimation des équations structurelles (7), (8) et (9) par la méthode de Double Moindres Carrés/DMC

NB : Seul Eviews est utilisé ici (avec plusieurs possibilités d'estimation).

► **Estimation des équations prises individuellement :** les commandes Eviews sont :

- * Equation eq7.TSLS C0t c Pt Pt(-1) Wt @ c t Gt WAt Tt Xt(-1) Pt(-1) Klag : pour Eq.(7);
- * Equation eq8.TSLS It c Pt Pt(-1) Klag @ c t Gt WAt Tt Xt(-1) Pt(-1) Klag : pour Eq.(8);
- * Equation eq9.TSLS WIt c Xt Xt(-1) t @ c t Gt WAt Tt Xt(-1) Pt(-1) Klag : pour Eq.(9).



____ Précisons que les variables/séries placées après “@” sont exogènes (elles constituent les instruments).

Dependent Variable: COT				
Method: Two-Stage Least Squares				
Date: 12/14/13 Time: 23:18				
Sample (adjusted): 1921 1941				
Included observations: 21 after adjustments				
Instrument list: C T GT WAT TT XT(-1) PT(-1) KLAG				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	16.55476	1.467979	11.27725	0.0000
PT	0.017302	0.131205	0.131872	0.8966
PT(-1)	0.216234	0.119222	1.813714	0.0874
WT	0.810183	0.044735	18.11069	0.0000
R-squared	0.976711	Mean dependent var	53.99524	
Adjusted R-squared	0.972601	S.D. dependent var	6.860866	
S.E. of regression	1.135659	Sum squared resid	21.92525	
F-statistic	225.9334	Durbin-Watson stat	1.485072	
Prob(F-statistic)	0.000000			

Dependent Variable: IT				
Method: Two-Stage Least Squares				
Date: 12/14/13 Time: 23:26				
Sample (adjusted): 1921 1941				
Included observations: 21 after adjustments				
Instrument list: C T GT WAT TT XT(-1) PT(-1) KLAG				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	20.27821	8.383249	2.418896	0.0271
PT	0.150222	0.192534	0.780237	0.4460
PT(-1)	0.615944	0.180926	3.404398	0.0034
KLAG	-0.157788	0.040152	-3.929751	0.0011
R-squared	0.884884	Mean dependent var	1.266667	
Adjusted R-squared	0.864569	S.D. dependent var	3.551948	
S.E. of regression	1.307149	Sum squared resid	29.04686	
F-statistic	41.20019	Durbin-Watson stat	2.085334	
Prob(F-statistic)	0.000000			

Dependent Variable: WIT				
Method: Two-Stage Least Squares				
Date: 12/14/13 Time: 23:28				
Sample (adjusted): 1921 1941				
Included observations: 21 after adjustments				
Instrument list: C T GT WAT TT XT(-1) PT(-1) KLAG				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	10.62799	3.076818	3.454217	0.0030
XT	0.438859	0.039603	11.08155	0.0000
XT(-1)	0.146674	0.043164	3.398063	0.0034
T	0.130396	0.032388	4.026001	0.0009
R-squared	0.987414	Mean dependent var	36.36190	
Adjusted R-squared	0.985193	S.D. dependent var	6.304401	
S.E. of regression	0.767155	Sum squared resid	10.00496	
F-statistic	424.1940	Durbin-Watson stat	1.963416	
Prob(F-statistic)	0.000000			



► Estimation des équations prises ensemble (en une seule instruction) :

Sur Eviews, suivre le chemin : *Object/New Object.../System* → ok → dans la boîte de dialogue, écrire :

$$\begin{cases} COT = C(1) + C(2)*PT + C(3)*PT(-1) + C(4)*WT \\ IT = C(5) + C(6)*PT + C(7)*PT(-1) + C(8)*KLAG \\ WIT = C(9) + C(10)*XT + C(11)*XT(-1) + C(12)*T \\ inst\ c\ t\ Gt\ WAt\ Tt\ Xt(-1)\ Pt(-1)\ Klag \end{cases}$$

Ensuite, cliquer sur **Estimate** → (dans « Method », sélectionner : *Two-Stage Least Squares*) → ok :

System: SYS1					
Estimation Method: Two-Stage Least Squares					
Sample: 1921 1941					
Included observations: 21					
Total system (balanced) observations 63					
	Coefficient	Std. Error	t-Statistic	Prob.	
C(1)	16.55476	1.467979	11.27725	0.0000	
C(2)	0.017302	0.131205	0.131872	0.8956	
C(3)	0.216234	0.119222	1.813714	0.0756	
C(4)	0.810183	0.044735	18.11069	0.0000	
C(5)	20.27821	8.383249	2.418896	0.0192	
C(6)	0.150222	0.192534	0.780237	0.4389	
C(7)	0.615944	0.180926	3.404398	0.0013	
C(8)	-0.157788	0.040152	-3.929751	0.0003	
C(9)	10.62799	3.076818	3.454217	0.0011	
C(10)	0.438859	0.039603	11.08155	0.0000	
C(11)	0.146674	0.043164	3.398063	0.0013	
C(12)	0.130396	0.032388	4.026001	0.0002	
Determinant residual covariance		0.287714			
Equation: COT = C(1) + C(2)*PT + C(3)*PT(-1) + C(4)*WT					
Instruments: C T GT WAT TT XT(-1) PT(-1) KLAG					
Observations: 21					
R-squared	0.976711	Mean dependent var		53.99524	
Adjusted R-squared	0.972601	S.D. dependent var		6.860866	
S.E. of regression	1.135659	Sum squared resid		21.92525	
Durbin-Watson stat	1.485072				
Equation: IT = C(5) + C(6)*PT + C(7)*PT(-1) + C(8)*KLAG					
Instruments: C T GT WAT TT XT(-1) PT(-1) KLAG					
Observations: 21					
R-squared	0.884884	Mean dependent var		1.266667	
Adjusted R-squared	0.864569	S.D. dependent var		3.551948	
S.E. of regression	1.307149	Sum squared resid		29.04686	
Durbin-Watson stat	2.085334				
Equation: WIT = C(9) + C(10)*XT + C(11)*XT(-1) + C(12)*T					



Instruments: C T GT WAT TT XT(-1) PT(-1) KLAG			
Observations: 21			
R-squared	0.987414	Mean dependent var	36.36190
Adjusted R-squared	0.985193	S.D. dependent var	6.304401
S.E. of regression	0.767155	Sum squared resid	10.00496
Durbin-Watson stat	1.963416		

► **Résolution du modèle de Klein sur Eviews :**

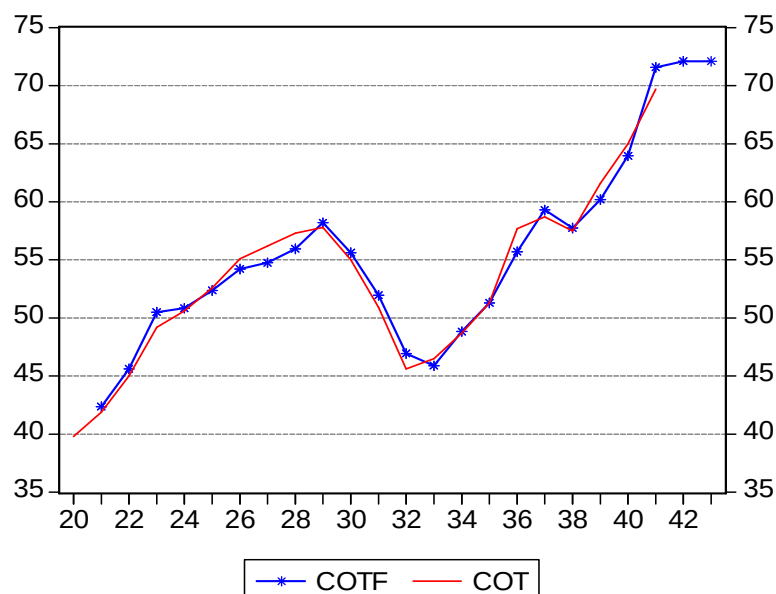
Sur Eviews, suivre : **Object/New Object.../Model** → OK → cliquer sur l'onglet **Text** → écrire : Assign @ all F et saisir les équations estimées comme suit :

$$\left\{ \begin{array}{l} COT_t = 16.55 + 0.017*Pt + 0.216*Pt(-1) + 0.810*Wt \\ It = 20.28 + 0.15*Pt + 0.616*Pt(-1) - 0.158*KLAG \\ WIt = 10.63 + 0.439*Xt + 0.147*Xt(-1) + 0.130*t \\ Xt = COT_t + It + Gt \\ Pt = Xt - WIt - Tt \\ Kt = It + Klag \end{array} \right.$$

Ensuite, cliquer sur l'onglet **Solve**.

► **Prévision** : dans l'output de chaque équation estimée, cliquer sur l'onglet **Forecast** (les variables se terminant par « f » sont ajustées ou prédites).

Exemple : **plot COTf COTf**



II. Simulations ou scénarios sur des modèles à équations simultanées

Cas pratique 4.4 : Modèle de KLEIN (1950) : Scénarios et/ou Simulations

- **Rappel sur le Modèle** : Soit le modèle macroéconomique de KLEIN (il s'agit d'un système linéaire à équations simultanées avec 6 équations et 11 variables économiques, en plus du vecteur « unité ») suivant :

$$\left\{ \begin{array}{l} CO_t = \pi_0 + \pi_1 P_t + \pi_2 P_{t-1} + \pi_3 (W_{it} + W_{at}) + e_{1t} \dots \dots (7) \\ I_t = \gamma_0 + \gamma_1 P_t + \gamma_2 P_{t-1} + \gamma_3 K_{t-1} + e_{2t} \dots \dots \dots (8) \\ W_{it} = \alpha_0 + \alpha_1 X_t + \alpha_2 X_{t-1} + \alpha_3 t + e_{3t} \dots \dots \dots (9) \dots \dots \dots [4.5] \\ X_t = CO_t + I_t + G_t \dots \dots \dots (10) \\ P_t = X_t - W_{it} - T_t \dots \dots \dots (11) \\ K_t = I_t + K_{t-1} \dots \dots \dots (12) \end{array} \right.$$

Avec :

- o CO_t : Consommation au temps t ;
- o P_t : Profits au temps t ;
- o W_{it} : Salaires (wage) versés par le secteur « industrie » au temps t ;
- o W_{at} : Salaires versés par le secteur « administration » au temps t ;
- o I_t : Investissement au temps t ;
- o X_t : Production industrielle au temps t ;
- o K_t : Stock de capital à la fin de l'année t ;
- o T_t : Impôt sur les bénéfices (profits) au temps t ;
- o G_t : Dépenses publiques au temps t ;
- o t : le temps (la tendance/trend)
- o $K_{t-1} = KLAG$.

► **Travail demandé :**

- Estimer les équations structurelles/de comportement par les DMC ;
- Trouver la solution initiale du modèle statique ou Résoudre le modèle (obtenir le Baseline ou solution de base) ;
- Analyser ou élaborer le 1^{er} **Scénario** : toutes choses restant égales par ailleurs/Ceteris paribus (c.à.d. bloquer le modèle : autres variables exogènes (G_t et T_t) restant constantes), si la masse salariale/« W_{at} » augmente de « 1 Milliard » (maintenue sur les 4 ans qui suivent), comment réagiront toutes les variables endogènes ? (autrement dit, quel est l'impact dans tout le système si « W_{at} » varie de 1 milliard positivement ?) ;
- Analyser le 2^{ème} **Scénario** : Ceteris paribus (bloquer le modèle), si les dépenses publiques non salariales « G_t » augmentent de « 1 milliard » (maintenue sur 4 ans suivants), comment réagira le système (tenir compte des canaux de transmission dans l'analyse) ?



- Analyser le 3^{ème} **Scénario** : « *Ceteris paribus* » (mais ne pas bloquer « t »), si les impôts « Tt » baissent de « 2 milliards » en 1942 et les dépenses publiques non salariales « Gt » augmentent de « 1,5 milliard » en même temps (maintenue jusqu'en 1945), comment réagira le système ?
- Analyser le 4^{ème} **Scénario** : Pour un objectif de « croissance annuelle à 15% » (Cfr Production industrielle « Xt ») sur 4 ans à venir, quelles mesures de politique budgétaire (« Gt », « Tt » ou « Wat ») mettre en place ?

1) Estimation des équations structurelles

Les résultats d'estimation des équations structurelles 7, 8 et 9 par les DMC ont été déjà présentés précédemment, seulement nous voulons ici insister sur la manière d'estimer l'équation 7 (relative à la consommation) qui accuse une variable composite : soit (Wat+Wit). Pour cette équation, procéder comme suit (obtenir les résultats suivants) :

```

Estimation Command:
=====
Equation eq7.TSLS COT C PT PT(-1) (WIT+WAT) @ C T GT WAT TT XT(-1) PT(-1) KLAG

Estimation Equation:
=====
COT = C(1) + C(2)*PT + C(3)*PT(-1) + C(4)*(WIT+WAT)

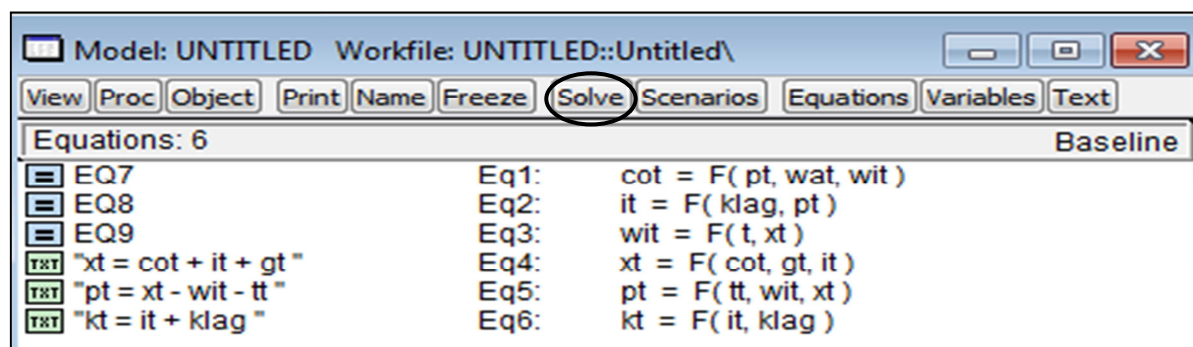
Substituted Coefficients:
=====
COT = 16.5547557654 + 0.0173022117998*PT + 0.216234040485*PT(-1) + 0.810182697599*(WIT+WAT)

```

2) Recherche de la solution initiale (baseline)

Dans **EViews 6** :

- o Sélectionner les équations enregistrées dans le Workfile→Clic droit→Open as Model→Ok→un espace s'ouvre (nous l'appelons « **Boîte A** ») ;
- o Dans l'espace « Boîte A », faire clic droit→insert (pour insérer les relations d'équilibre : 10, 11 et 12) →ok. A ce niveau, la « boîte A » s'affiche comme suit :



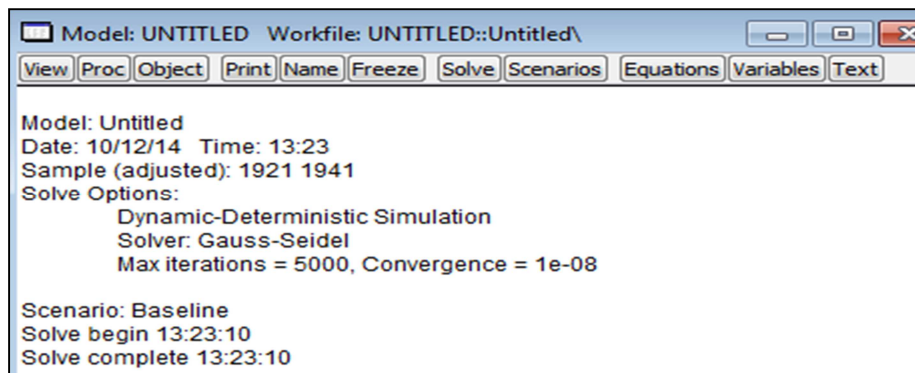
- o Dans la « Boîte A » ci-dessus, cliquer sur **Solve** pour résoudre le modèle (simulation statique/solution mathématique ou solution de base/Baseline)→dans la boîte de dialogue qui s'affichera, faire le choix sur



le type de simulation et l'algorithme de résolution. Dans EViews (Cfr onglets : « Basic Options » et « Solver ») et en théorie, l'on distingue :

- * 2 types de simulations : simulations déterministe et stochastique ;
- * 2 types de solution : solutions statique et dynamique ;
- * 3 types d'algorithme : algorithme de Gauss-Seidel (populaire/courant), celui de Newton et celui de Broyden.

Dans notre cas, nous optons pour la simulation dynamique déterministe et retenons l'algorithme de Gauss-Seidel → cliquer sur Ok pour valider. Ainsi les infos suivantes s'affichent (range : 1920 – 1941) :

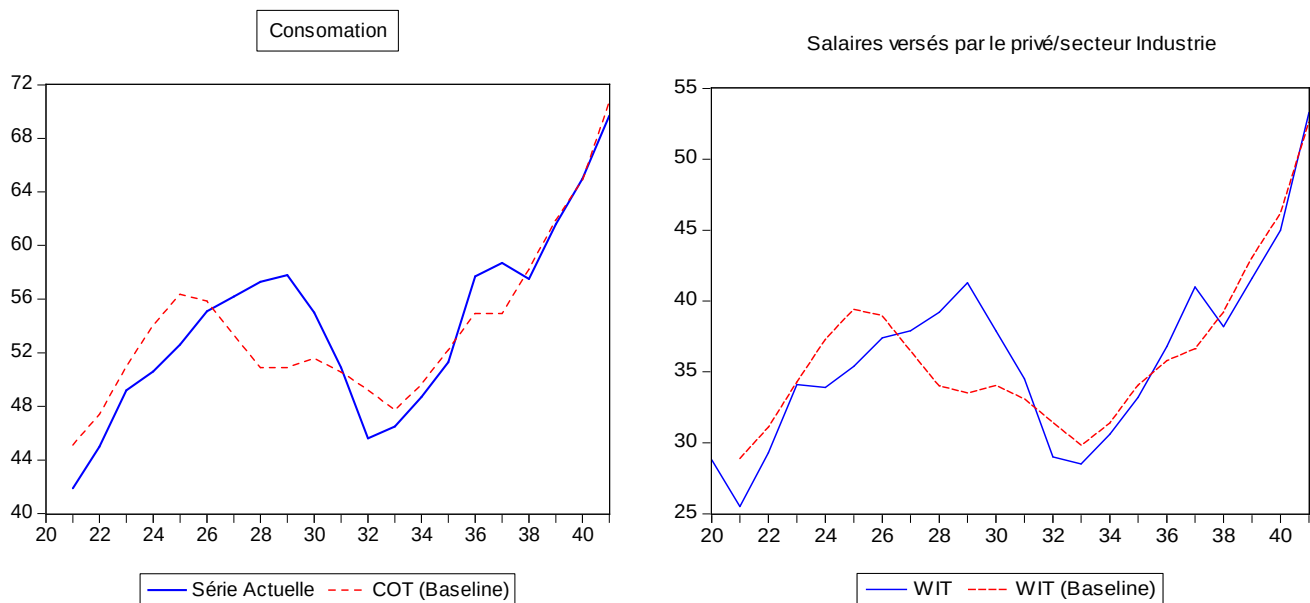
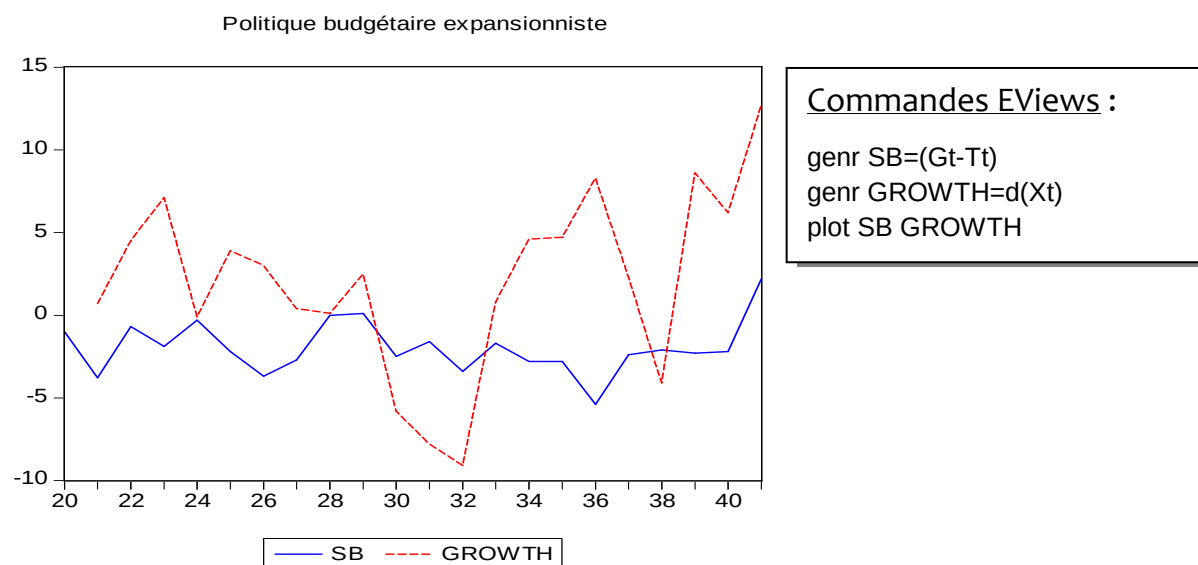


Il s'agit entre autres du nombre d'itérations, de l'algorithme et type de simulation, et des dates de début et de fin de résolution. Ces informations nous aident à vérifier que le modèle a passé le test de convergence et que la solution mathématique du modèle a été trouvée (c'est la procédure de validation du modèle). Toutefois, Cfr d'autres critères pour évaluer la qualité du modèle résolu/prédit, à savoir :

- RMSE : la Racine de l'Erreur Quadratique Moyenne ;
- MAE : l'Erreur Absolue Moyenne ;
- Theil Inequality Coefficient (coefficient d'inégalité de theil (u)).
- L'on peut se servir aussi des graphiques (plot série brut et série simulée).

NB : simulation de base = simulation historique (possibilité d'évaluer et quantifier l'impact des politiques alternatives). Servons-nous de graphes (simulation historique: elle permet d'évaluer/quantifier les politiques économiques alternatives/simulations) :



(i) Consommation/Cot et Salaires du secteur industrie/Wit : actual vs baseline(ii) Solde budgétaire/SB et Taux de croissance économique/GROWTH : actual vs baseline

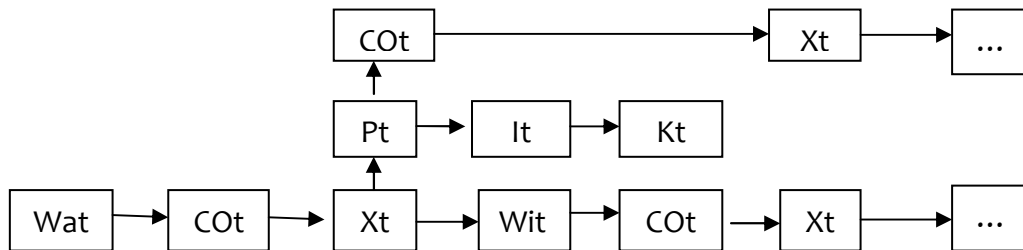
Note : considérant l'évolution du solde budgétaire, il va sans dire que la politique budgétaire poursuivie par les USA (entre 1921 et 1941) a été expansionniste (*accroissement des dépenses traduit par des déficits budgétaires persistants*), avec des taux de croissance économique globalement positifs comme corollaire dans une certaine mesure (*ceteris paribus*).



3) Scénarios et projections/simulations

(i) Scénario 1 : Ceteris paribus, « Wat » ↑ de 1 milliard sur 4 ans à venir

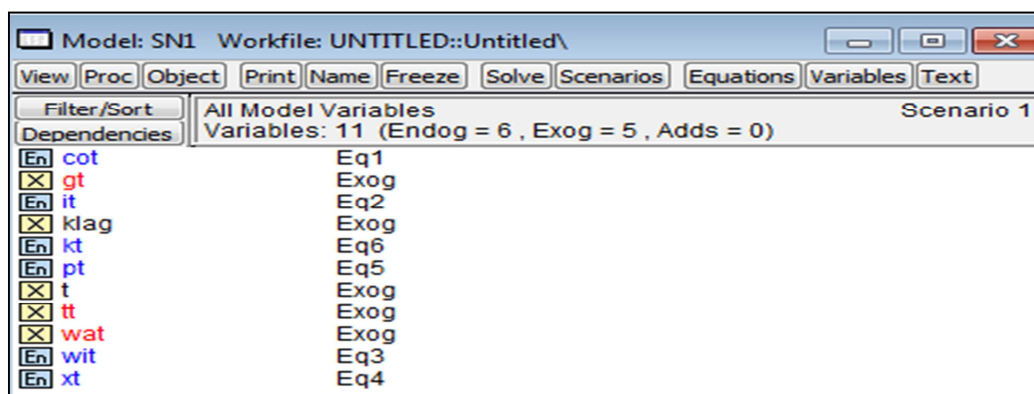
Schéma : Selon le modèle :



_____ **Note :** Bouger une variable exogène (Wat), ceteris paribus signifie ignorer les effets d'autres variables exogènes (Gt, Tt, t et Klag) ou bloquer les autres variables exogènes (Gt, Tt, t et Klag *considérées constantes dont les dérivées sont nulles : mêmes valeurs entre 1942 et 1945*).

Etapes à suivre :

► **Créer un scénario** : dans « **Boîte A** », procéder comme suit (suivre) : View/Scenarios... → **Scenario 1** → dans « scenario Overrides », préciser les variables exogènes du modèle : Gt, Wat et Tt → cliquer sur Ok. Ensuite, dans « **Boîte A** », cliquer sur l'onglet « **Variables** » pour vérifier que l'on ne s'est pas trompé dans la déclaration des variables exogènes. Les informations suivantes s'affichent (en rouge, les variables exogènes déclarées ; en noire celles qui sont automatiquement considérées comme exogènes parce que ne constituant pas des équations ; et, en bleu les variables endogènes/équations du modèle) :



► Préciser les valeurs prises par les 3 variables exogènes (G_t , W_{at} et T_t) et la période de projection. Pour ce faire :

- Etendre l'échantillon (période historique) jusqu'en 1945 (période de projection): *dans le Workfile (EViews), double-cliquer sur « Range »→End date: 1945→Ok*;
- Créer des nouvelles variables exogènes associées à chaque scénario (pour éviter de supprimer les valeurs de départ/initiales). Ex: Clic droit sur « Wat » dans la « **Boîte A** » ci-dessus→Propriétés→Cocher « use override



series in scenario→Yes→Ok : ainsi, l'on a créé « Wat_1 » par exemple (en faire autant pour « Gt », « Tt », « Klag » et « t »). Ensuite, ouvrir « Wat_1 » et compléter les valeurs relatives à la période de projection comme suit :

- * 1942 : valeur de 1941 + 1 milliard ;
- * 1943...1945 : valeur de 1942 recopiée.

_____ **Note** : l'on peut suivre aussi « Quick/Generate Series » pour procéder de la même façon.

- Reproduire les mêmes valeurs pour les autres variables exogènes « Gt et Tt », mais aussi « Klag et t » (ceteris paribus).
- **Simuler le modèle** : Revenir à la « Boîte A », cliquer sur l'onglet « **Solve** »→[Active : **Scenario 1**]→Ok →le message suivant s'affiche (indiquant que la simulation a réussi) :

```
Model: SN1
Date: 10/12/14 Time: 22:35
Sample (adjusted): 1921 1945
Solve Options:
    Dynamic-Deterministic Simulation
    Solver: Gauss-Seidel
    Max iterations = 5000, Convergence = 1e-08

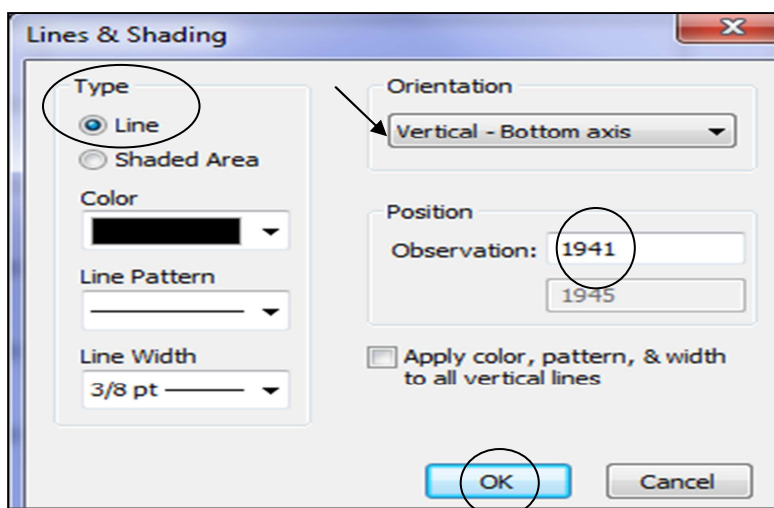
Scenario: Scenario 1
Solve begin 22:35:34
Solve complete 22:35:34
```

Et nous pouvons observer ce qui suit (Cfr données sur Workfile) :

Période	COt_1	Xt_1	Wit_1	Kt_1	Pt_1	Wat_1	It_1
1941	70.71087	88.23387	52.58769	208.2230	24.04618	8.5	3.722996
1942	77.73379	98.55536	58.99777	211.5216	27.95760	9.5	7.021570
1943	82.58550	106.2363	63.88252	214.3508	30.75381	9.5	9.850832
1944	86.10797	111.7791	67.44162	216.3711	32.73749	9.5	11.87113
1945	88.62365	115.7273	69.98730	217.8037	34.14001	9.5	13.30366

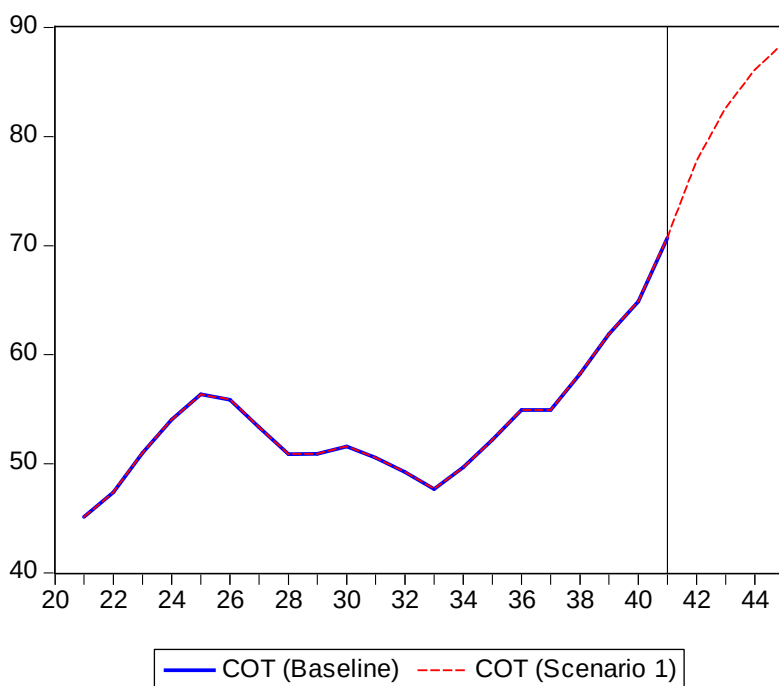
- Graphiques (plot série brut et série simulée) : dans la boîte de dialogue du graphique obtenu (concernant la consommation), cliquer sur l'onglet « Freeze »→Line/Shade→La boîte ci-dessous complète la procédure (pour tracer une ligne verticale à partir d'une date, 1941 par exemple) :





Cliquer sur « Ok » pour obtenir le graphique suivant :

[Consommation simulée \(scénario 1\)](#)

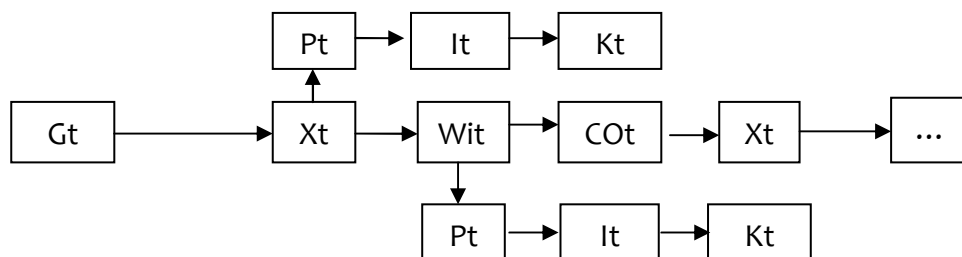


Commentaires : L'augmentation de la masse salariale de l'administration publique/Wat d'1 milliard, ceteris paribus, a des effets positifs sur tout le système (consommation, production, etc.) durant les 4 années à venir : les variables endogènes subissent des accroissements à des rythmes de plus en plus ralentit d'une année à une autre (Cfr données et graphique ci-haut).

(ii) **Scénario 2 : Ceteris paribus, « Gt » ↑ de 1 milliard sur 4 ans à venir**

Schéma : Selon le modèle :





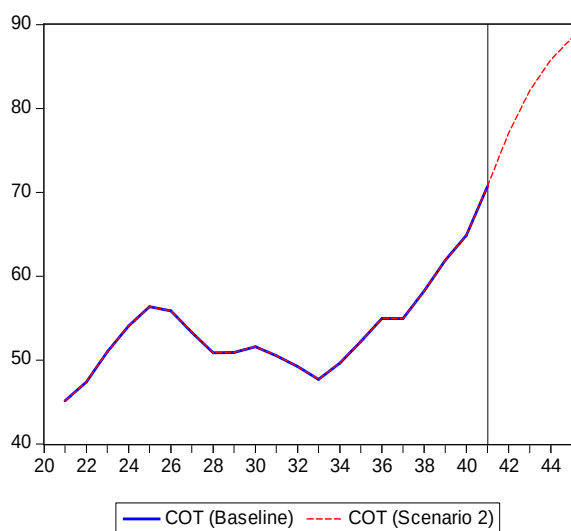
Simulation du modèle : Les résultats intermédiaires s'affichent comme suit :

Model: SN1 Date: 10/12/14 Time: 23:48 Sample (adjusted): 1921 1945 Solve Options: Dynamic-Deterministic Simulation Solver: Gauss-Seidel Max iterations = 5000, Convergence = 1e-08 Scenario: Scenario 2 Solve begin: 23:48:19 Solve complete: 23:48:19	Filter/Sort All Model Variables Scenario 2 Variables: 11 (Endog = 6, Exog = 5, Adds = 0) Dependencies En cot Eq1 X gt Exog En it Eq2 X klag Exog En kt Eq6 En pt Eq5 X t Exog X tt Exog X wat Exog En wit Eq3 En xt Eq4
---	---

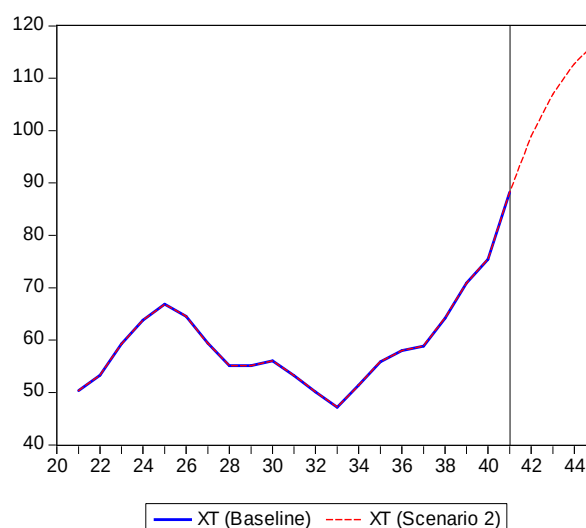
Nous pouvons observer également ce qui suit (Cfr données sur Workfile) :

Période	COt_2	Xt_2	Wit_2	Kt_2	Pt_2	Gt_2	It_2
1941	70.71087	88.23387	52.58769	208.223	24.04618	13.80000	3.722996
1942	77.04957	98.90021	59.14911	211.5506	28.15110	14.80000	7.050639
1943	82.11165	106.9328	64.23875	214.5211	31.09404	14.80000	10.02113
1944	85.80976	112.7576	67.97318	216.6478	33.18441	14.80000	12.14783
1945	88.45786	116.9154	70.65222	218.1575	34.66317	14.80000	13.65752

Consommation simulée (Scénario 2)



Production simulée (Scénario 2)



(iii) Scénario 3 : Ceteris paribus, « Wat » ↑ de 1,5 milliard et « Tt » ↓ de 2 milliards sur 4 ans à venir (« t » évolue)

Simulation du modèle : Le modèle a passé le test de convergence avec succès suivant les informations ci-dessous :

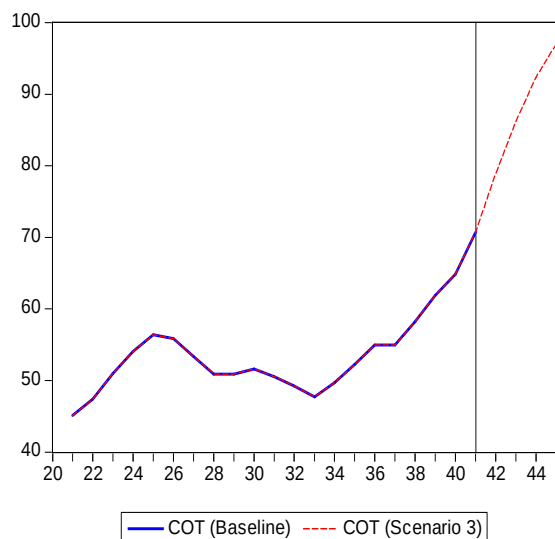
```
Model: SN1
Date: 10/13/14 Time: 00:49
Sample (adjusted): 1921 1945
Solve Options:
    Dynamic-Deterministic Simulation
    Solver: Gauss-Seidel
    Max iterations = 5000, Convergence = 1e-08

Scenario: Scenario 3
Solve begin 00:49:54
Solve complete 00:49:54
```

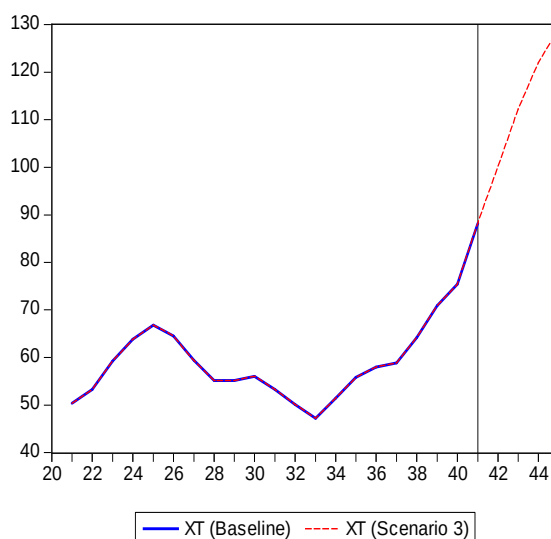
Nous pouvons observer également ce qui suit (Cfr données sur Workfile) :

Période	COT_3	Xt_3	Wit_3	Kt_3	Pt_3	Tt_3	It_3	Wat_3
1941	70.71087	88.23387	52.58769	208.2230	24.04618	11.60000	3.722996	8.500000
1942	78.82363	100.0522	59.78508	211.9286	30.66716	9.600000	7.428605	10.00000
1943	86.18423	112.2382	66.99685	216.7540	35.64137	9.600000	12.25399	10.00000
1944	92.34510	121.9975	73.19755	220.3524	39.19994	9.600000	15.85239	10.00000
1945	96.98410	129.2010	77.92070	222.9169	41.68028	9.600000	18.41688	10.00000

Consommation simulée (Scénario 3)



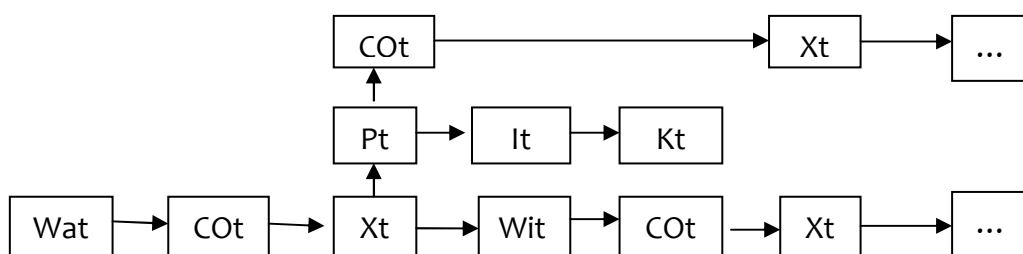
Production simulée (Scénario 3)



(iv) Scénario 4 : Pour que « X_t » ↑ de 15% chaque année (sur 4 ans suivants : de 1942 à 1945), comment « Wat » doit évoluer durant ces 4 ans ?

Note : Il est pour nous question de simuler l'évolution requise des variables budgétaires (G_t , W_t et T_t) correspondant à la trajectoire imposée à la production industrielle (X_t) ou croissance économique. L'on notera que nous considérons « Wat » comme variable de contrôle traduisant la politique budgétaire ; car, elle nous permet justement de boucler le modèle⁽¹⁾, toutes choses restantes égales par ailleurs (« T_t », « G_t », « t » et « Klag » considérées constantes). En pratique, l'exercice consiste à bloquer les variables exogènes « T_t , G_t , t et Klag » (sauf « Wat »), ouvrir la production industrielle « X_t » et obtenir (simulation) l'évolution de « Wat » compatible avec l'objectif fixé (accroissement de « X_t » de 15% chaque année, sur les 4 ans suivants).

Schéma : Selon le modèle (rappelons) :



Etapes à suivre :

- **Créer un scénario :** dans « **Boîte A** », procéder comme suit (suivre) : View/Scenarios... → Create New Scenario → **Scenario 4** → dans « scenario Overrides », préciser les variables exogènes du modèle : G_t , W_t et T_t → Ok.
- **Bloquer les variables exogènes** « T_t , G_t , t et Klag » (sauf « Wat ») – c.à.d., pour ces 4 variables, répéter les valeurs en 1941 sur les années suivantes – et ouvrir la production industrielle « X_t ». Pour ce faire :
 - Générer des nouvelles variables exogènes associées au 4^{ème} scénario créé. Ex : Clic droit sur « G_t » dans la « **Boîte A** » → Properties → Cocher « use override series in scenario » → Yes → Ok : ainsi, l'on a créé « G_{t_4} » par exemple (en faire autant pour « T_t », « W_t », « Klag » et « t »). Ensuite, ouvrir les variables générées et compléter les valeurs relatives à la période de projection (répéter les mêmes valeurs, sauf pour « Wat » qui gardera des données manquantes).
 - L'objectif de 15% du taux de croissance économique impose à la production industrielle (« X_t ») une trajectoire de croissance. D'où, pour la période « 1942–1945 », ouvrir « X_t » en suivant : Quik/Generate Series... → Dans la boîte de dialogue qui s'affiche, saisir à tour de rôle

¹ La variable qui permet de boucler le modèle est celle d'où part l'impact et où finissent les effets/impacts. « Wat » est ainsi choisie au détriment de « T_t » et « G_t ».



(cliquer sur « ok » et reprendre le processus pour donner d'autres instructions) :

- { Xaim=Xt → sample : 1920 1945 → ok : *définir le 4^{ème} Scénario*
- { Xaim=Xaim(-1)*1.15 → sample : 1942 1945 → ok : *Xt ↑ 15% entre 1942 et 1945 (imposer une trajectoire à « Xt » ou ouvrir « Xt »)*

En passant, dans « **Boîte A** », cliquer sur l'onglet « **Variables** » pour observer les variables exogènes définies/identifiées dans le système (en rouge) :

Filter/Sort	All Model Variables	Scenario 4
Dependencies	Variables: 11 (Endog = 6 , Exog = 5 , Adds = 0)	
☒ cot	Eq1	
☒ gt	Exog	
☒ it	Eq2	
☒ klag	Exog	
☒ kt	Eq6	
☒ pt	Eq5	
☒ t	Exog	
☒ tt	Exog	
☒ wat	Exog	
☒ wit	Eq3	
☒ xt	Eq4	

► **Simulation** : Dans « **Boîte A** », suivre : Proc→Solve Control for Target... → taper « **en gras** » (dans chaque case) :

- o Solve for values of Control Variable : **Wat**
- o so model solves the Target Variable : **Xt**
- o to the values of the Trajectory Series : **Xaim**
- o Sample : **1942 1945** → Ok.

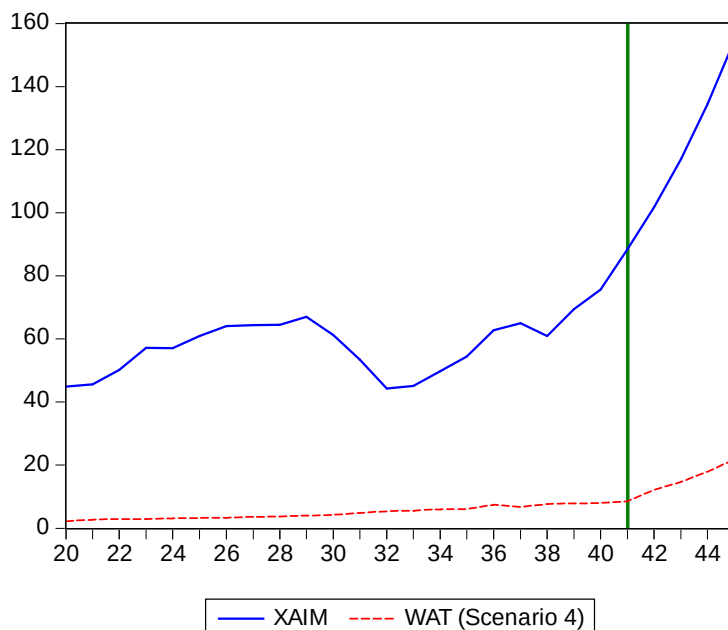
NB : Xt « variable cible/target » ; Xaim « sentier de la variable cible sur la période de projection », et Wat « variable de contrôle ». Après avoir cliqué sur « Ok », nous obtenons l'évolution des salaires publics « Wat » (entre 1942 et 1945 : période de projection) – compatible avec le taux de croissance de la production industrielle « Xt » fixé à 15% pour chaque année – comme suit :

Model: SN1	
Date: 10/19/14 Time: 14:22	
Control:	WAT
Target:	XT
Trajectory:	XAIM
Sample:	1942 1945
Solving for 'Scenario 4 WAT' over 1942-1945, to set 'Scenario 4 XT' = XAIM	
=====	
	WAT_4
=====	
1942	12.15098
1943	14.62541
1944	17.97404
1945	21.82495
=====	



Graphiques (plot variable cible « Xaim » vs variable de contrôle « Wat ») :

Niveau des salaires publics compatibles avec l'objectif d'une croissance économique à 15% (entre 1942 et 1945) aux USA

**Exportation des « résultats/simulations/projections » d'EViews vers Excel :**

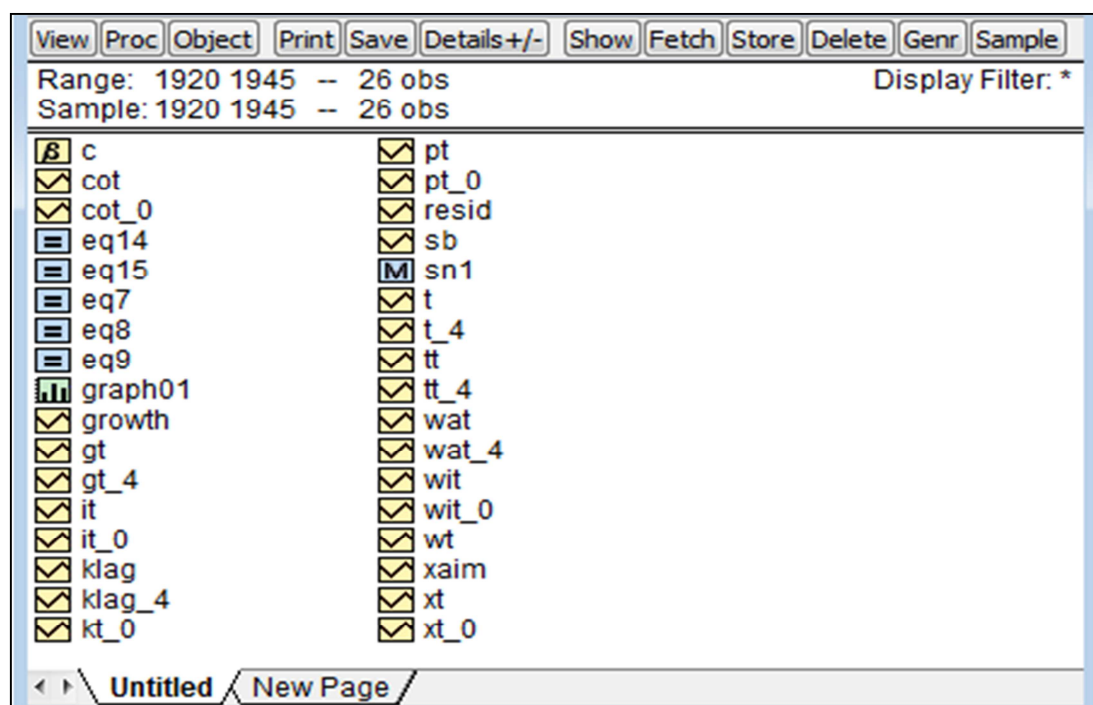
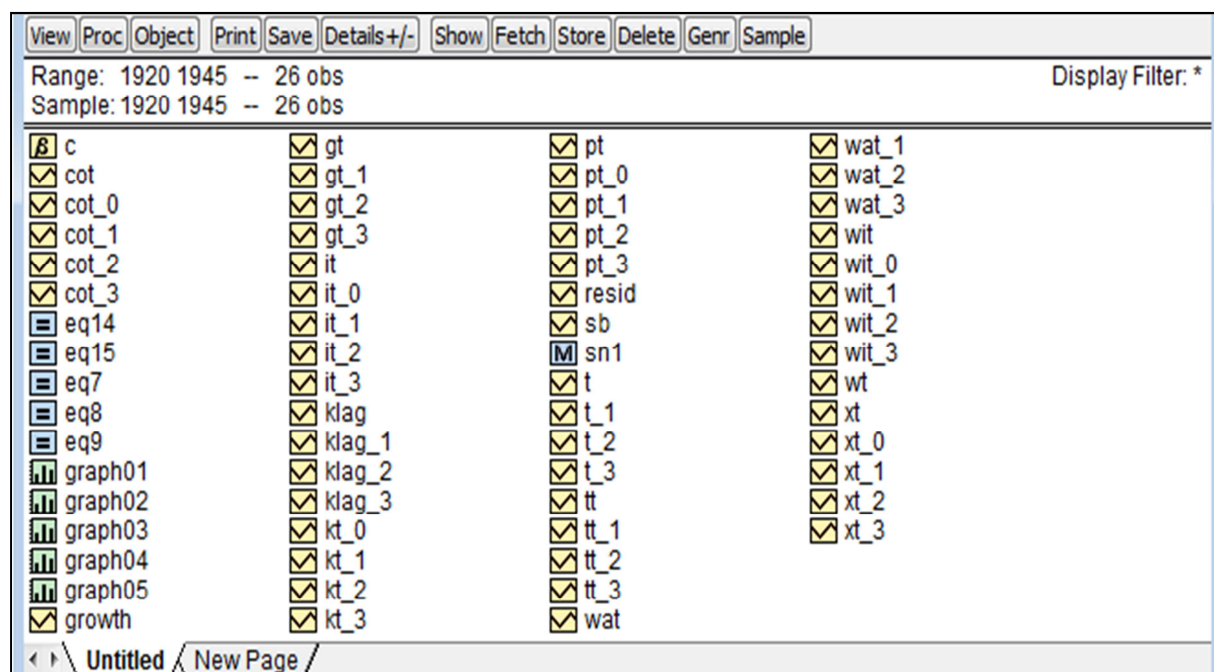
Dans EViews (6), sélectionner les variables à transférer et suivre : *File/Export/Write Text-Lotus-Excel* → Choisir le nom du fichier Excel « *simulation result* » et son emplacement, et préciser le type de fichier de sortie (« Excel ») → dans « *Series to export* », vérifier que toutes les variables sélectionnées apparaissent dans la boîte de dialogue qui s'affiche → Ok.

Mise à jour du modèle :

- Plusieurs raisons justifient la mise à jour d'un modèle macro-économétrique, entre autres :
 - o Le renouvellement des données par *an/trimestre/mois* : cela implique ré-estimation avec l'ajout des nouvelles données ;
 - o Les mutations économiques → l'augmentation des équations de comportement → prise en compte des chocs ou nouveaux enjeux économiques. L'on notera que, pour les modèles annuels par exemple, ils sont mis à jour au moins une fois par an.
- Étapes à suivre : dans EViews, après avoir introduit toutes les nouvelles données quantitatives et/ou qualitatives, suivre (Cfr « Boîte A ») : *Proc/Links/Update All Links-Recompile Model*. Après, résoudre/simuler/projeter les politiques économiques ou sociales.



Annexe : Workfiles produits (Cfr équations simultanées : Modèle de Klein – Scénarios et/ou Simulations).



Cas pratique 4.5 : Estimation d'un modèle à équations simultanées par la méthode de DMC et Simulation/Scénarios

1) **Modèle**

Soit le modèle du multiplicateur/accélérateur d'un pays donné :

$$\begin{cases} C_t = b_0 + b_1 Y_t + e_{1t} \\ I_t = a_0 + a_1 Y_{t-1} + a_2 Y_t + e_{2t} \dots \dots [4.6] \\ Y_t = C_t + I_t + G_t \end{cases}$$

Avec : a_2 = accélérateur et G_t = dépenses gouvernementales.

NB : multiplicateur (mesure de l'effet d'une variation de l'investissement/ I_t sur le revenu/ Y_t) et accélérateur (mesure l'effet sur l'investissement de l'augmentation du revenu ou de la consommation/ C_t).

2) **Identification** : dans le système, $G=3$ et $K=2$.

Equations	Identification	Conclusion	Méthode
1 ^{ère} équation	$g'=2$ (Ct et Yt) ; $k'=0$ et $r=0$. $(G-g')+(K-k')+r ? (g'-1)$ $(3-2)+(2-0)+0 > 2-1$ $3 > 1$	La 1 ^{ère} équation est sur-identifiée	Doubles Moindres Carrés (DMC)
2 ^{ème} équation	$g'=2$ (It et Yt) ; $k'=1$ (Yt-1) et $r=0$. $(G-g')+(K-k')+r ? (g'-1)$ $(3-2)+(2-1)+0 > 2-1$ $2 > 1$	La 2 ^{ème} équation est sur-identifiée	Doubles Moindres Carrés/DMC

3) **Estimation** :

Nous estimons ces deux équations (séparément et on les enregistre dans le workfile) par la méthode de doubles moindres carrés/DMC. Sur EViews (6), taper :

```
create a 1967 1987
data Ct Gt It Yt
equation eq1.tsls Ct c Yt @ Yt(-1) Gt
equation eq2.tsls It c Yt Yt(-1) @ Yt(-1) Gt
```

Dependent Variable: CT Method: Two-Stage Least Squares Date: 09/07/14 Time: 13:10 Sample (adjusted): 1968 1987 Included observations: 20 after adjustments Instrument list: YT(-1) GT					Dependent Variable: IT Method: Two-Stage Least Squares Date: 09/07/14 Time: 13:10 Sample (adjusted): 1968 1987 Included observations: 20 after adjustments Instrument list: YT(-1) GT				
Variable	Coefficient	Std. Error	t-Statistic	Prob.	Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	252.7037	1072.934	0.235526	0.8165	C	-3178.841	3083.016	-1.031082	0.3169
YT	0.783710	0.004367	179.4807	0.0000	YT	0.478260	0.078715	6.075804	0.0000
					YT(-1)	-0.351815	0.126089	-2.790207	0.0126
R-squared	0.999446	Mean dependent var	90969.04		R-squared	0.961953	Mean dependent var	26592.52	
Adjusted R-squared	0.999416	S.D. dependent var	175079.7		Adjusted R-squared	0.957477	S.D. dependent var	58882.71	
S.E. of regression	4232.542	Sum squared resid	3.22E+08		S.E. of regression	12142.27	Sum squared resid	2.51E+09	
F-statistic	32213.32	Durbin-Watson stat	1.982982		F-statistic	221.6616	Durbin-Watson stat	1.472678	
Prob(F-statistic)	0.000000	Second-Stage SSR	5.32E+09		Prob(F-statistic)	0.000000	Second-Stage SSR	5.15E+08	



4) Simulations/Scénarios

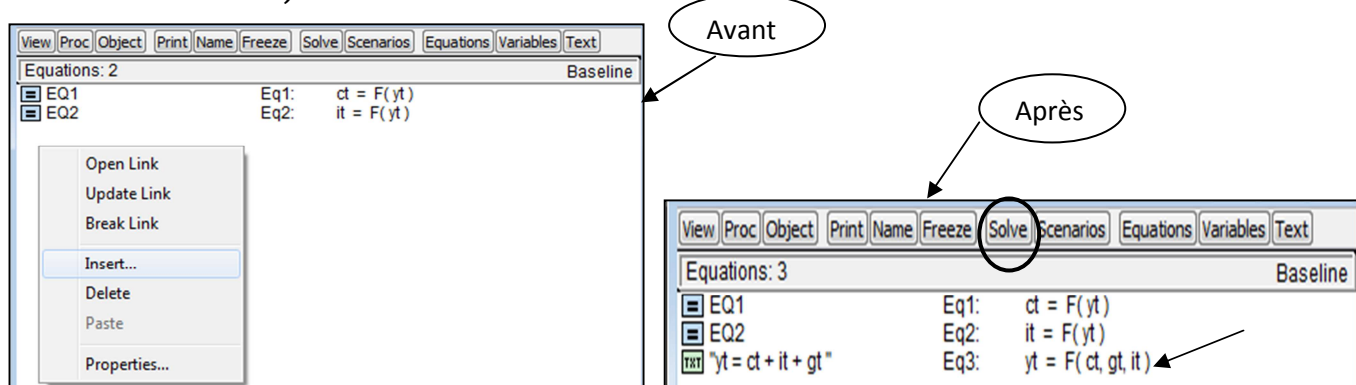
Travail demandé (les questions sont ordonnées selon les étapes à suivre pour une simulation) :

- Trouver la solution initiale du modèle statique ou Résoudre le modèle (obtenir le *Baseline* ou solution de base) ;
- Analyser ou élaborer le 1^{er} Scénario : Si « G_t » augmente de « 1% », comment réagiront C_t et I_t ? (autrement dit, quel est l'impact dans tout le système si G_t varie de 1% positivement ?) ;
- Analyser le 2^{ème} Scénario : Si « Y_{t-1} » augmente de « 2% », toutes choses restant égales par ailleurs (c.à.d. bloquer le modèle), comment réagira le système (tenir compte des canaux de transmission dans l'analyse) ?
- Analyser le 3^{ème} Scénario : Si « G_t » augmente de « 1% », lever l'hypothèse « *ceteris paribus* » (c.à.d. ne pas bloquer le modèle), et que « Y_{t-1} » augmente en même temps de 1,5%, comment réagira le système ?
- Analyser le 4^{ème} Scénario : Si « G_t » baisse de « 3% », toutes choses restant égales par ailleurs (c.à.d. bloquer le modèle), comment réagira le système ? Pour ce scénario, envisager une simulation hors échantillon (**sur 1 an, c.à.d. « 1988 »**).

a) Recherche de la solution initiale (Baseline)

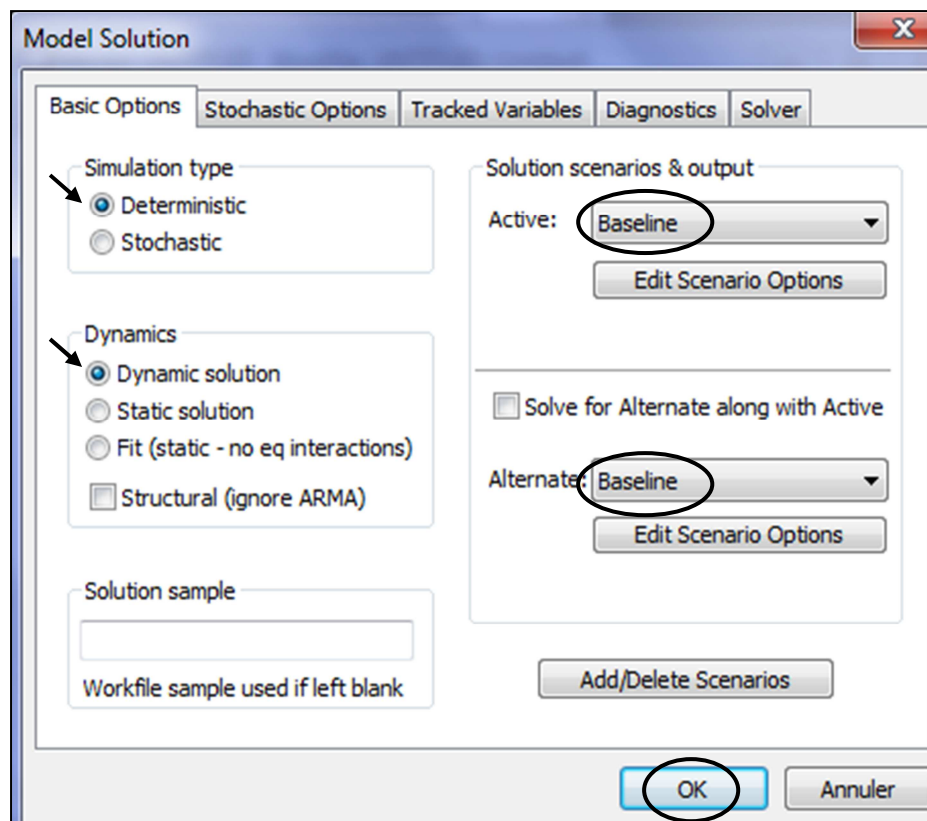
Après avoir estimé nos deux équations séparément et les avoir enregistré dans le Workfile/fichier de travail, procéder comme suit :

- o Sélectionner les équations enregistrées dans le Workfile → Clic droit → Open as Model → Ok → un espace s'ouvre (nous l'appelons « **Boîte A** ») ;
- o Dans l'espace « Boîte A », faire clic droit → insert (pour insérer la relation d'équilibre) : $Y_t = C_t + I_t + G_t$ → ok (la boîte de dialogue A ci-dessous est illustrative) :

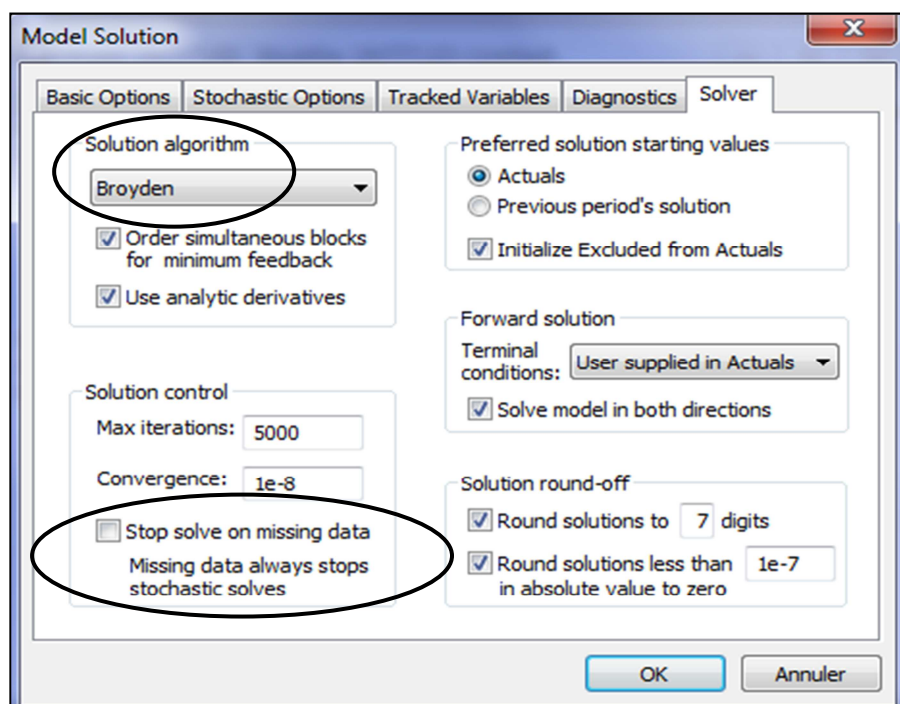


- o Dans la boîte de dialogue ci-haut (à droite), cliquer sur « **Solve** » pour résoudre le modèle (simulation statique ou solution de base/Baseline) → la boîte de dialogue ci-dessous apparaît :





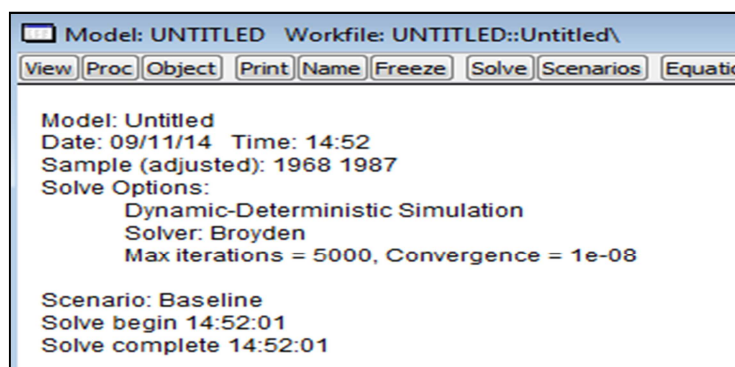
Cliquer sur l'onglet « **Solver** » pour préciser ce qui suit (à décocher) :



Ensuite, cliquer sur « **Ok** » pour obtenir la solution de base. En fait, les solutions de base sont des variables générées par EViews avec des indices « **_o** » ; dans notre cas, ces variables sont : **Ct_o**, **It_o** et **Yt_o**. L'on notera que le modèle est soluble (équations *juste-identifiées* ou *sur-identifiées*) s'il passe le test de convergence avec succès (*assurance pour la solution mathématique du modèle*). Donc, le message



suivant devra s'afficher (il indique le nombre d'itérations, le type de simulation et l'algorithme utilisés, aussi les dates de début et de fin d'itérations/résolution) :



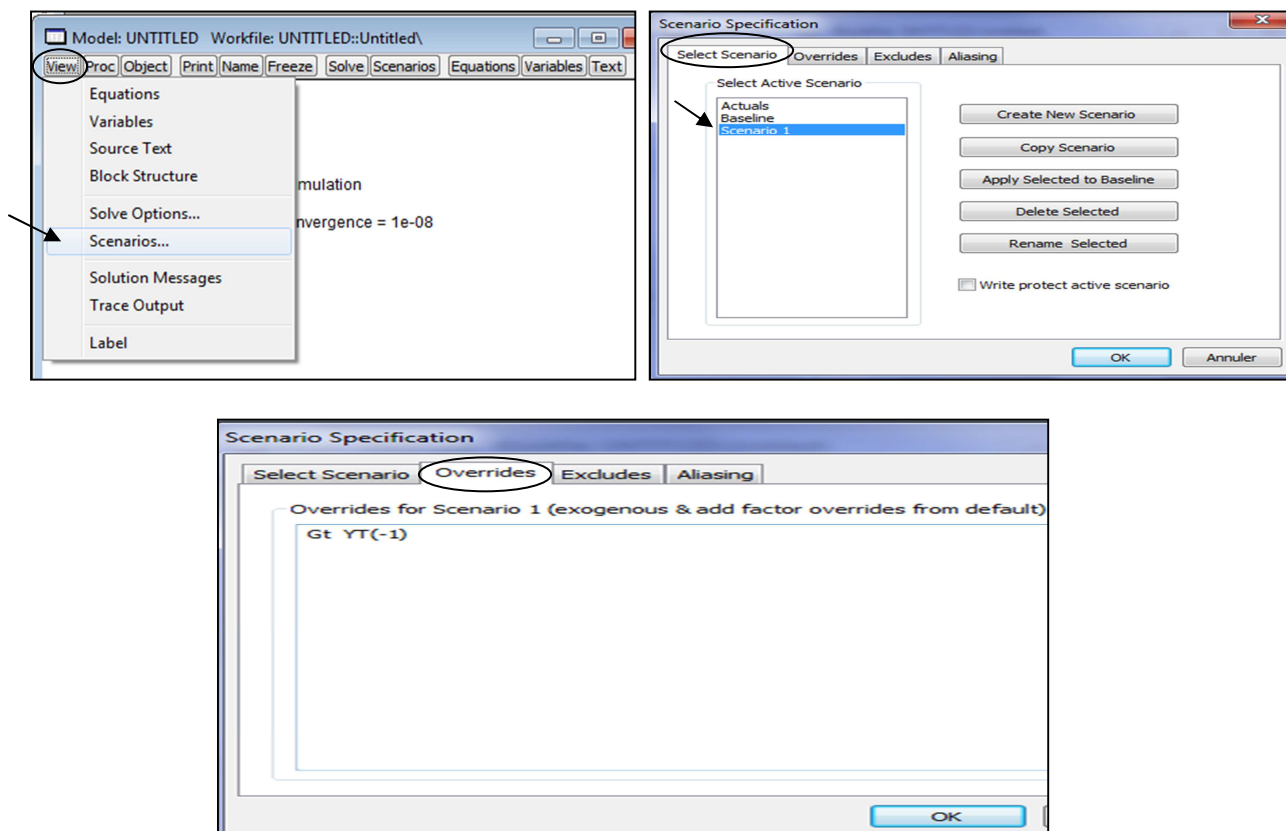
b) Elaboration du 1^{er} Scénario et analyses: si « Gt augmente de 1%, ceteris paribus »

_____ **Schéma** : Selon le modèle, $G_t \rightarrow Y_t \rightarrow C_t$ et $G_t \rightarrow Y_t \rightarrow I_t$

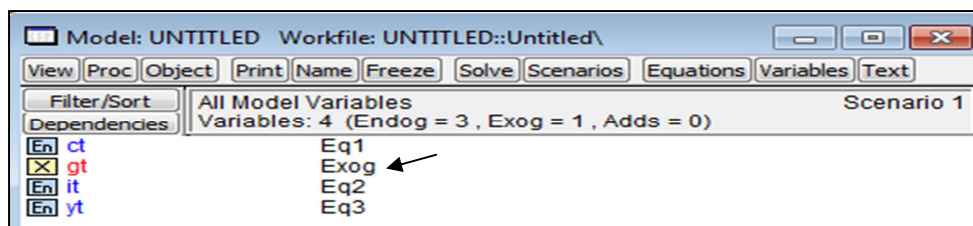
_____ **Note** : Bouger une variable exogène (G_t), ceteris paribus signifie ignorer les effets d'autres variables exogènes (Y_{t-1}) ou bloquer les autres variables exogènes (Y_{t-1} considérée constante dont la dérivée est nulle).

_____ **Etapes à suivre** :

- Dans « **Boîte A** », procéder comme suit (suivre) : View/Scenarios... → **Scenario 1** → dans « scenario Overrides », préciser les variables exogènes du modèle : G_t et Y_{t-1} . Les figures ci-dessous sont illustratives :



- Dans « **Boîte A** », cliquer sur l'onglet « **Variables** » pour vérifier que l'on ne s'est pas trompé dans la déclaration des variables exogènes. Les informations suivantes s'affichent :



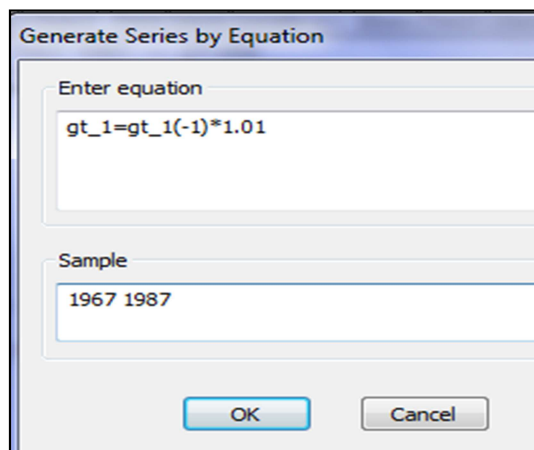
En rouge, les variables exogènes déclarées, et en bleu, les variables endogènes du modèle. Dans notre cas, « Gt » seulement apparaît comme exogène, alors que « Yt-1 » a été également déclarée exogène. *L'on notera ainsi que les variables décalées n'apparaissent pas dans la boîte A, à moins de les générer sous un autre nom à des fins de simulation.*

► **Simulation (Gt ↑ 1% en 1987) :**

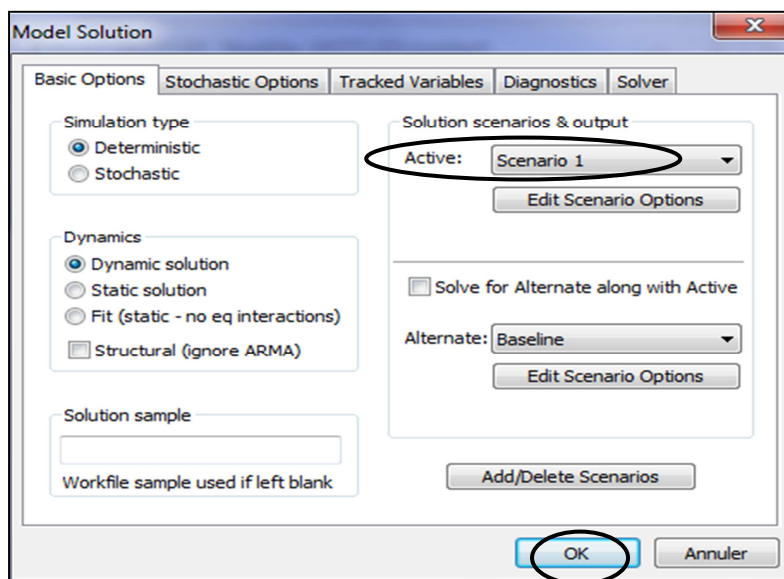
- o La simulation est similaire à la prévision dans l'échantillon (bloquer est nécessaire, mais l'on peut ouvrir/lever l'hypothèse ceteris paribus) comme en dehors de l'échantillon (blocage automatique, soit toujours bloquer). *Ici, nous faisons une simulation dans l'échantillon (1987) ;*
- o L'on devra bouger « Gt » et bloquer « Yt-1 » pour voir la réaction du modèle (toutes les variables endogènes) ;
- o Deux possibilités pour simuler : rester sur la barre de commande ou aller dans : *Quik/Generate Series...* → Dans la boîte de dialogue qui s'affiche, saisir à tour de rôle (cliquer sur « ok » et reprendre le processus pour donner d'autres instructions), avec « Gt_1 » traduisant la variable du scénario 1 :

$gt_1 = gt \rightarrow \text{sample : 1967 1987} \rightarrow \text{ok : définir le 1}^{er} \text{ Scénario}$
 $gt_1 = gt_1(-1) * 1.01 \rightarrow \text{sample : 1987 1987} \rightarrow \text{ok : Gt } \uparrow 1\% \text{ en 1987 (ouvrir « Gt »)}$
 $ylag_1 = yt(-1) \rightarrow \text{sample : 1967 1987} \rightarrow \text{ok : définir le 1}^{er} \text{ Scénario (bloquer « Yt-1 »)}$
 $ylag_1 = @elem(ylag_1, 1986) \rightarrow \text{sample : 1987 1987} \rightarrow \text{ok : bloquer « Yt-1 » (données 1986=1987 : Constance).}$

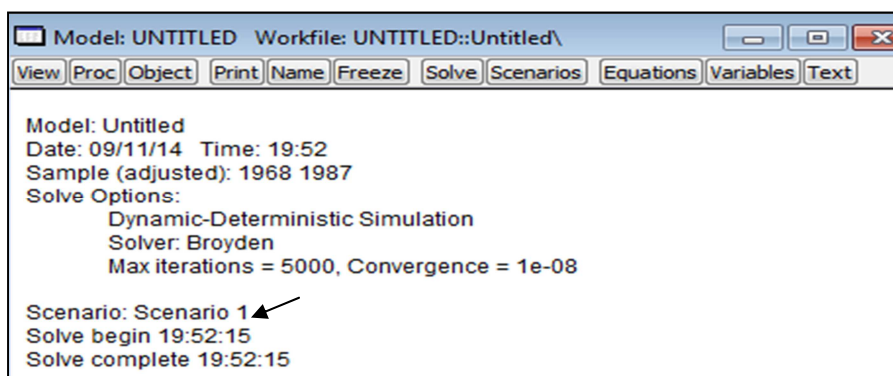
A titre illustratif, observons la boîte de dialogue ci-dessous :



- o Revenir à la « Boîte A », cliquer sur l'onglet « **Solve** » → [Active : **Scenario 1**], la boîte de dialogue ci-après complète la procédure¹ :



_____ **Résultats** : Après avoir cliqué sur « Ok » (Cfr figure ci-dessus), le message suivant s'affiche (indiquant que la simulation a réussi)



Et nous pouvons observer ce qui suit (Cfr données sur Workfile) :

Période	Gt_1	Yt_1	It_1	Ct_1
1986	38.600	7.398.695	1.561.413	5.798.683
1987	38.986	9.798.528	2.080.087	7.679.455
%	1%	32.4%	33.2%	32.4%

_____ **Analyses** :

¹ La simulation est différente du Scénario du fait que celui-ci cherche à saisir l'évolution d'une variable compatible avec une trajectoire imposée à d'autres variables.



c) **Elaboration du 2^{ème} Scénario et Analyses : « Yt-1 » augmente de « 2% », ceteris paribus**

Schéma : Selon le modèle, $Y_{t-1} \rightarrow I_t \rightarrow Y_t \rightarrow C_t \rightarrow Y_t \rightarrow I_t$

Note : Cette fois ci, on va bouger Yt-1 et bloquer Gt.

Etapes à suivre :

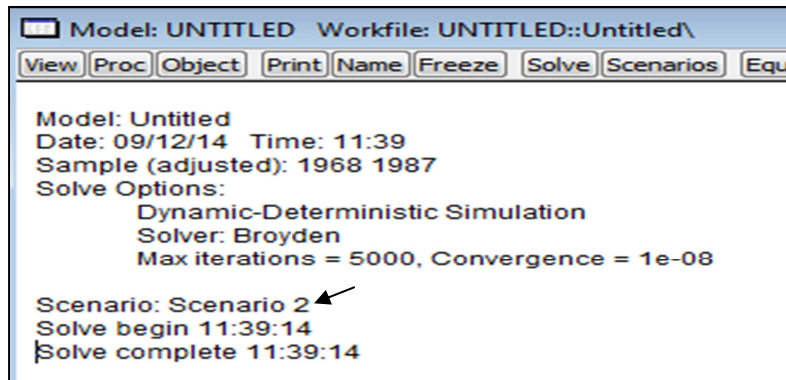
- ▶ Dans « **Boîte A** », pour créer un nouveau scénario, procéder comme suit (suivre) : View/Scenarios... → Create New Scenario → Sélectionner « **Scenario 2** ».

▶ **Simulation (Yt-1 ↑ 2% en 1987) :**

- o Ici, nous faisons une simulation dans l'échantillon (1987) ;
- o L'on devra bouger « Yt-1 » et bloquer « Gt » pour voir la réaction du modèle (toutes les variables endogènes) ;
- o Suivre : Quik/Generate Series... → Dans la boîte de dialogue qui s'affiche, saisir à tour de rôle, avec « YLAG_2 » traduisant la variable du scénario 2 :

$ylag_2 = yt(-1) \rightarrow$ sample : 1967 1987 → ok : définir le 2^{ème} Scénario
 $ylag_2 = ylag_2(-1) * 1.02 \rightarrow$ sample : 1987 1987 → ok : Yt-1 ↑ 2% en 1987
 $gt_2 = gt \rightarrow$ sample : 1967 1987 → ok : bloquer « Gt »
 $gt_2 = @elem(gt_2, 1986) \rightarrow$ sample : 1987 1987 → ok : bloquer « Gt » pour 1987.

- o Revenir à la « Boîte A », cliquer sur l'onglet « **Solve** » → [Active : **Scenario 2**] → Ok, le message suivant s'affiche (indiquant que la simulation a réussi) :



Et nous pouvons observer ce qui suit (Cfr données sur Workfile) :

Période	Ylag_2	Yt_2	It_2	Ct_2
1986	358.800	7.398.695	1.561.413	5.798.683
1987	365.976	9.621.355	1.995.352	7.540.603
%	2%	30.04%	27.79%	30.04%

Analyses :



d) Elaboration du 3^{ème} Scénario et Analyses : « Gt » augmente de « 1% » et « Yt-1 » augmente en même temps de 1,5% (en 1987)

Schéma : Pour rappel, selon le modèle : $Y_{t-1} \rightarrow I_t \rightarrow Y_t \rightarrow C_t$ et $G_t \rightarrow Y_t \rightarrow I_t$

Note : Cette fois ci, on va bouger Y_{t-1} et G_t en même temps.

Etapes à suivre :

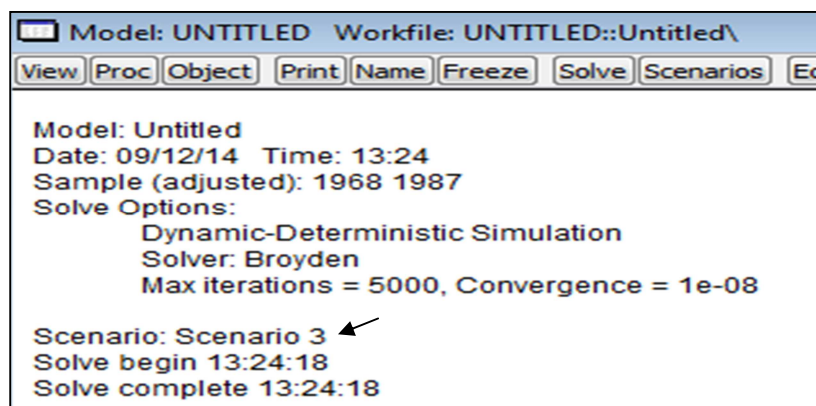
► Dans « **Boîte A** », procéder comme suit (suivre) : View/Scenarios... → Create New Scenario → Sélectionner « **Scenario 3** ».

► **Simulation ($Y_{t-1} \uparrow 1,5\%$ et $G_t \uparrow 1\%$ en 1987) :**

- o Ici, nous faisons une simulation dans l'échantillon (1987) ;
- o L'on devra bouger, en même temps, « Y_{t-1} » et « G_t » pour voir la réaction du modèle (toutes les variables endogènes) ;
- o Suivre : Quik/Generate Series... → Dans la boîte de dialogue qui s'affiche, saisir à tour de rôle, avec « $YLAG_3$ » et « Gt_3 » traduisant les variables du scénario 3 :

$\left\{ \begin{array}{l} ylag_3 = yt(-1) \rightarrow \text{sample : 1967 1987} \rightarrow \text{ok : définir le 3^{ème} Scénario sur « Yt-1 »} \\ ylag_3 = ylag_3(-1) * 1.015 \rightarrow \text{sample : 1987 1987} \rightarrow \text{ok : } Y_{t-1} \uparrow 1,5\% \text{ en 1987} \\ gt_3 = gt \rightarrow \text{sample : 1967 1987} \rightarrow \text{ok : définir le 3^{ème} Scénario sur « Gt »} \\ gt_3 = gt(-1) * 1.01 \rightarrow \text{sample : 1987 1987} \rightarrow \text{ok : } G_t \uparrow 1\% \text{ en 1987.} \end{array} \right.$

- o Revenir à la « Boîte A », cliquer sur l'onglet « **Solve** » → [Active : **Scenario 3**] → Ok, le message suivant s'affiche (indiquant que la simulation a réussi) :



Et nous pouvons observer ce qui suit (Cfr données sur Workfile) :

Période	Gt_3	Ylag_3	Yt_3	It_3	Ct_3
1986	38.600	358.800	7.398.695	1.561.413	5.798.683
1987	38.986	364.182	9.621.355	1.995.352	7.540.603
%	1%	1.5%	30.04%	27.79%	30.04%

Analyses :



e) **Elaboration du 4^{ème} Scénario et Analyses : « Gt » baisse de « 3% », ceteris paribus, mais prévision/simulation hors échantillon (en 1988)**

Schéma : Pour rappel, selon le modèle : $G_t \rightarrow Y_t \rightarrow I_t$ et $G_t \rightarrow Y_t \rightarrow C_t$

Note : on bouge « Gt », « Yt-1 » reste constante (bloquée).

Étapes à suivre :

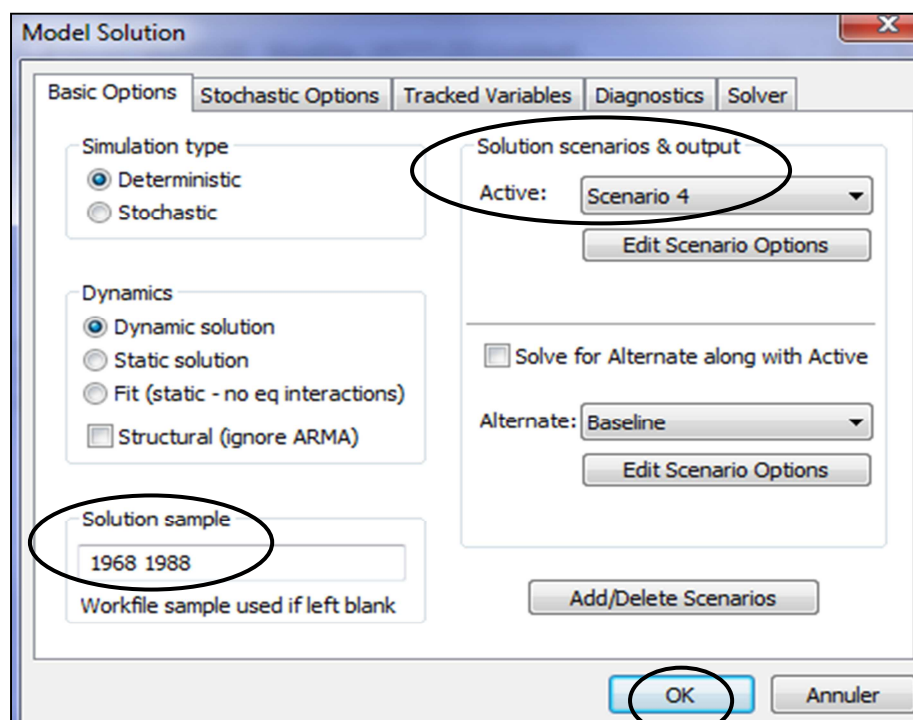
► Dans « **Boîte A** », procéder comme suit (suivre) : View/Scenarios... → Create New Scenario → Sélectionner « **Scenario 4** ».

► **Simulation (Gt ↓ 3% en 1988, ceteris paribus) :**

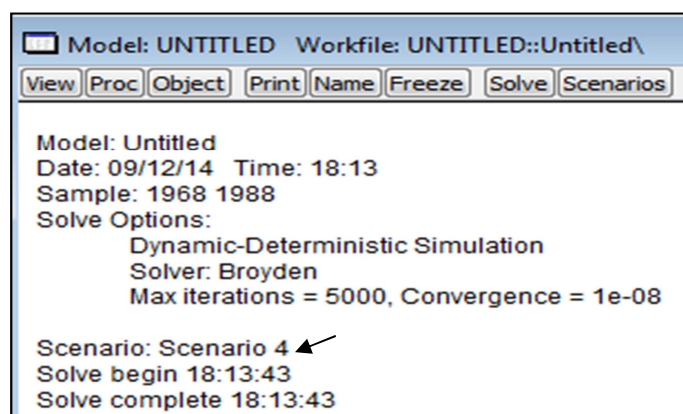
- o Ici, nous faisons une simulation **hors** échantillon (1988) ;
- o L'on devra bouger seulement « Gt » pour voir la réaction du modèle (toutes les variables endogènes) en dehors de l'échantillon ;
- o Suivre :
 - o Cfr Workfile : Proc → Structure/Resize Current Page... → dans « End date », indiquer « 1988 » → Ok → Yes : Elargir l'horizon temporel jusqu'à 1988 ;
 - o Cfr Barre des menus : Quik/Generate Series... → Dans la boîte de dialogue qui s'affiche, saisir à tour de rôle, avec « Gt_4 » traduisant la variable du scénario 4 :

$ylag_4 = yt(-1) \rightarrow$ sample : 1967 **1988** → ok : bloquer « Yt-1 »
 $ylag_4 = @elem(ylag_4, 1987) \rightarrow$ sample : **1988 1988** → ok : bloquer "Yt-1" en **1988**
 $gt_4 = gt \rightarrow$ sample : 1967 **1988** → ok : définir le 4^{ème} Scénario
 $gt_4 = gt_4(-1) * 0.97 \rightarrow$ sample : **1988 1988** → ok : Gt ↓ 3% en **1988**

- o Revenir à la « **Boîte A** », cliquer sur l'onglet « **Solve** » → [Active : **Scenario 4**] → ... la boîte de dialogue ci-dessous complète la procédure :



→ Cliquer sur « Ok », le message suivant s'affiche (indiquant que la simulation a réussi) :

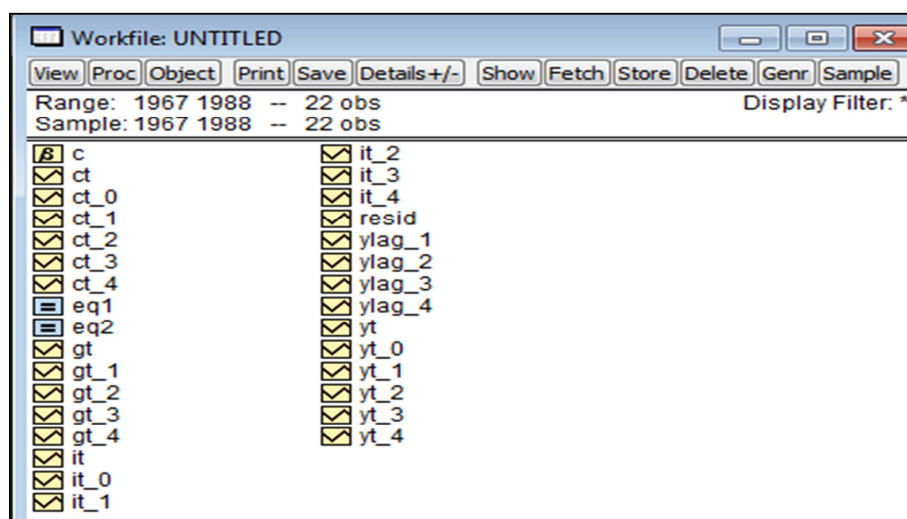


Et nous pouvons observer ce qui suit (Cfr données sur Workfile) :

Période	Gt_4	Yt_4	It_4	Ct_4
1987	85.400	9.621.355	1.995.352	7.540.603
1988	82.838	12.616.090	2.645.645	9.887.602
%	-3%	31.13%	32.59%	31.12%

Analyses :

Annexe : Paysage relatif à notre Workfile





CHAP V

MODELES DE PANEL

Cas pratique 5.1 : Investissement fonction du profit

► **Modèles :**

❖ Modèle de départ (général) : **Modèle 0**

$$I_{it} = c + a_i + b_1 P_{it} + e_{it} \dots \dots [5.1]$$

($1 \leq i \leq N$; $1 \leq t \leq T$; $N = 3$ firmes et $T = 10$ années (de 1982 à 1991, les années 1992 à 1995 sont à prévoir).

Avec :

- o I_{it} : Dépenses d'investissement pour la firme i au temps t ;
- o P_{it} : Profits réalisés par la firme i au temps t ;

❖ Modèle sans effets spécifiques et aléatoires : **Modèle 1**

$$I_{it} = c + b_1 P_{it} + e_{it} \dots \dots [5.2]$$

Sous l'hypothèse que : $a_1 = a_2 = a_3 = \dots = a_n = 0$ (2 paramètres à estimer).

❖ Modèle à effets fixes : **Modèle 2**

$$I_{it} = a_i + b_1 P_{it} + e_{it} \dots \dots [5.3]$$

Sous l'hypothèse que les coefficients sont identiques entre les firmes (sauf pour la constante). NB : 4 paramètres sont à estimer (avec « a_i » : les paramètres fixes non aléatoires).

❖ Modèle à effets aléatoires : **Modèle 3**

$$I_{it} = \mu + b_1 P_{it} + a_i + e_{it} \dots \dots [5.4]$$

Posons : $e_{it} = a_i + v_{it}$, ainsi « 5.4 » s'écrit :

$$I_{it} = \mu + b_1 P_{it} + e_{it} \dots \dots [5.4]'$$

Avec : a_i et v_{it} , respectivement l'effet spécifique et l'effet résiduel. NB : 5 paramètres sont à estimer ici.

► **Travail demandé :**

- Estimer le modèle à effets fixes (Modèle 2), le modèle à effets aléatoires (Modèle 3) et le modèle sans effets (Modèle 1) ;
- Effectuer les différents tests de spécification sur données de panel (test de Fisher/Likelihood Ratio Test ; test de Breusch et Pagan ; et test de Hausman) ;
- Effectuer les différents tests de stationnarité/racine unitaire sur données de panel (test de Im-Pesaran-Shin et celui de Levin-lin) et ceux de cointégration sur données de panel (test de Pedroni) ;
- Partant des modèles estimés (Modèle 2 et Modèle 3), faire une prévision à horizon de 4 ans ;



- Estimer un modèle à correction d'erreur à effets fixes ;
- Estimer un panel dynamique ;
- Effectuer les tests post-estimation sur panel (inférence Statistique) ;
- Estimer un modèle de panel à effets between et un modèle à effets aléatoires avec l'estimateur ml (maximum de vraisemblance) ;
- Estimer un panel corrigé d'hypothèses violées.

a) Estimation du modèle à effets fixes (Modèle 2), du modèle à effets aléatoires (Modèle 3) et du modèle sans effets

Extrait des données affichées sur Eviews (à gauche) et sur Stata (à droite) :

FIRME1-1983	26.30000	25.69000					
FIRME1-1984	2.620000	5.480000					
FIRME1-1985	14.94000	13.79000					
FIRME1-1986	15.80000	15.41000					
FIRME1-1987	12.20000	12.59000					
FIRME1-1988	14.93000	16.64000					
FIRME1-1989	29.82000	26.45000					
FIRME1-1990	20.32000	19.64000					
FIRME1-1991	4.770000	5.430000					
FIRME1-1992	NA	7.000000					
FIRME1-1993	NA	10.00000					
FIRME1-1994	NA	12.00000					
FIRME1-1995	NA	10.00000					
FIRME2-1982	20.30000	22.93000					
FIRME2-1983	17.47000	17.96000					
FIRME2-1984	9.310000	9.160000					
FIRME2-1985	18.01000	18.73000					
FIRME2-1986	7.630000	11.31000					
FIRME2-1987	19.84000	21.15000					

	фирме	annee	I	P
1	1	1982	13.32	12.85
2	1	1983	26.3	25.69
3	1	1984	2.62	5.48
4	1	1985	14.94	13.79
5	1	1986	15.8	15.41
6	1	1987	12.2	12.59
7	1	1988	14.93	16.64
8	1	1989	29.82	26.45
9	1	1990	20.32	19.64
10	1	1991	4.77	5.43
11	1	1992	.	7
12	1	1993	.	10
13	1	1994	.	12
14	1	1995	.	10
15	2	1982	20.3	22.93
16	2	1983	17.47	17.96
17	2	1984	9.31	9.16
18	2	1985	18.01	18.73
19	2	1986	7.63	11.31

► **Estimation du modèle à effets fixes (Modèle 2)**

Sur Stata, après avoir déclaré les données en panel à stata comme suit (le message en bas confirme la déclaration) : **tsset firme annee**

panel variable: firme, 1 to 3
time variable: annee, 1982 to 1995

faire : **xtreg I P, fe**

Fixed-effects (within) regression				Number of obs	=	30
Group variable (i): firme				Number of groups	=	3
R-sq: within	=	0.9483	Obs per group: min	=	10	
between	=	0.8629	avg	=	10.0	
overall	=	0.9247	max	=	10	
				F(1, 26)	=	477.26
corr(u_i, Xb) = -0.2295				Prob > F	=	0.0000

I		Coef.	Std. Err.	t	P> t	[95% Conf. Interval]

P		1.096293	.0501823	21.85	0.000	.9931422 1.199445
_cons		-1.361473	.8144482	-1.67	0.107	-3.035596 .312649

sigma_u		1.3851987				
sigma_e		1.728003				
rho		.39120601	(fraction of variance due to u_i)			

F test that all u_i=0:		F(2, 26) =	6.09	Prob > F = 0.0068		



Sur Eviews, faire : create a 1982 1995 , ensuite suivre le chemin :

Object/New Object.../Pool → (écrire les noms des firmes en colonne, précédés chacun d'une sous barre) → Sheet → (saisir les noms des variables suivi d'un « ? » chacun) → (copier et coller les observations) → Estimate → (cross section=Fixed, period=none) → (Method : LS..) → (Options : Ordinary) → ok.

Dependent Variable: I?				
Method: Pooled Least Squares				
Date: 12/15/13 Time: 18:05				
Sample (adjusted): 1982 1991				
Included observations: 10 after adjustments				
Cross-sections included: 3				
Total pool (balanced) observations: 30				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.361474	0.814448	-1.671651	0.1066
P?	1.096293	0.050182	21.84621	0.0000
Fixed Effects (Cross)				
_FIRME1-C	-0.016157			
_FIRME2-C	-1.377050			
_FIRME3-C	1.393207			
Effects Specification				
Cross-section fixed (dummy variables)				
R-squared	0.948693	Mean dependent var	15.04200	
Adjusted R-squared	0.942773	S.D. dependent var	7.223449	
S.E. of regression	1.728003	Akaike info criterion	4.055376	
Sum squared resid	77.63586	Schwarz criterion	4.242202	
Log likelihood	-56.83064	F-statistic	160.2517	
Durbin-Watson stat	1.834669	Prob(F-statistic)	0.000000	

Dans l'output, cliquer sur : View/Representations :

```

I_FIRME1 = -0.0161567914 - 1.361473563 + 1.096293457*P_FIRME1
I_FIRME2 = -1.377049788 - 1.361473563 + 1.096293457*P_FIRME2
I_FIRME3 = 1.39320658 - 1.361473563 + 1.096293457*P_FIRME3

```

► Estimation du modèle à effets aléatoires (Modèle 3)

Sur Stata, faire : `xtreg I P, re`

Random-effects GLS regression		Number of obs	=	30		
Group variable (i): firme		Number of groups	=	3		
R-sq: within	= 0.9483	Obs per group: min	=	10		
between	= 0.8629	avg	=	10.0		
overall	= 0.9247	max	=	10		
Random effects u_i ~ Gaussian		Wald chi2(1)	=	343.69		
corr(u_i, X) = 0 (assumed)		Prob > chi2	=	0.0000		

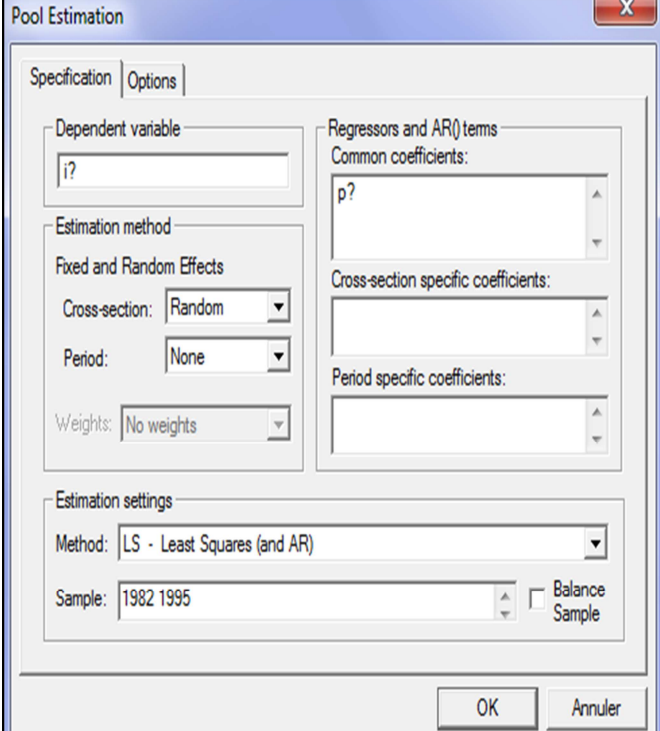
I		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]

P		1.056158	.0569703	18.54	0.000	.944498 1.167817
_cons		-.7609357	.92862	-0.82	0.413	-2.580998 1.059126

sigma_u		0				
sigma_e		1.728003				
rho		0	(fraction of variance due to u_i)			



Sur Eviews, suivre : *Object/New Object.../Pool* → (écrire les noms des individus en colonne, précédés chacun d'une sous barre) → *Sheet* → (saisir les noms des variables suivi d'un « ? » chacun) → (copier et coller les observations) → **Estimate** → (cross section=**random**, period=**none**) → (Method : **LS.**) → (Options : **Ordinary**) → ok.



Dependent Variable: I?

Method: Pooled EGLS (Cross-section random effects)

Date: 12/15/13 Time: 19:05

Sample (adjusted): 1982 1991

Included observations: 10 after adjustments

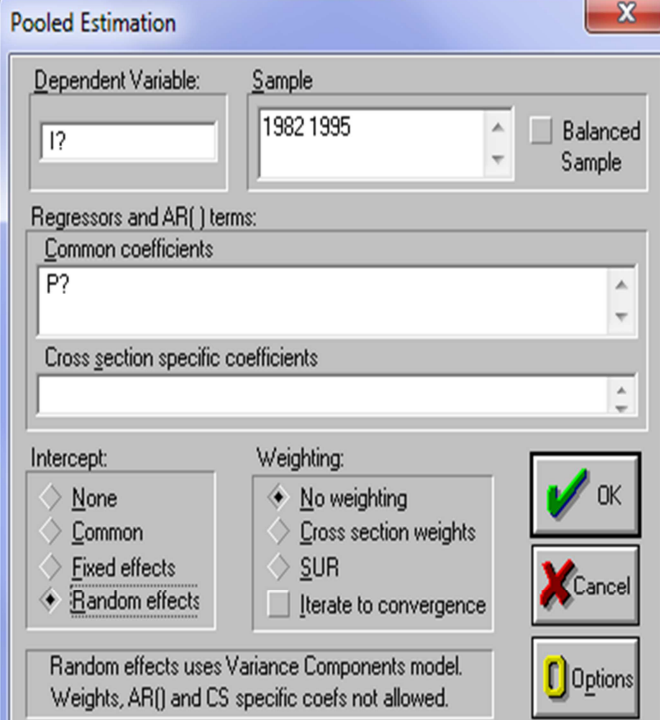
Cross-sections included: 3

Total pool (balanced) observations: 30

Swamy and Arora estimator of component variances

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.760936	0.795292	-0.956800	0.3469
P?	1.056158	0.048791	21.64671	0.0000
Random Effects (Cross)				
_FIRME1--C	0.000000			
_FIRME2--C	0.000000			
_FIRME3--C	0.000000			
Effects Specification				
Cross-section random S.D. / Rho			0.000000	0.0000
Idiosyncratic random S.D. / Rho			1.728003	1.0000
Weighted Statistics				
R-squared	0.924667	Mean dependent var	15.04200	
Adjusted R-squared	0.921977	S.D. dependent var	7.223449	
S.E. of regression	2.017697	Sum squared resid	113.9908	
F-statistic	343.6853	Durbin-Watson stat	1.250819	
Prob(F-statistic)	0.000000			
Unweighted Statistics				
R-squared	0.924667	Mean dependent var	15.04200	
Sum squared resid	113.9908	Durbin-Watson stat	1.250819	

Les résultats (figures) précédents sont produits sur Eviews 5. Sur Eviews 3, les résultats (figures) changent comme illustré ci-dessous :



Dependent Variable: I?

Method: GLS (Variance Components)

Date: 12/15/13 Time: 19:00

Sample: 1982 1991

Included observations: 10

Number of cross-sections used: 3

Total panel (balanced) observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.179589	0.991966	-0.181043	0.8576
P?	1.017305	0.063358	16.05645	0.0000
Random Effects				
_FIRME1--C	-0.019690			
_FIRME2--C	1.357057			
_FIRME3--C	-1.337367			
GLS Transformed Regression				
R-squared	0.901299	Mean dependent var	15.04200	
Adjusted R-squared	0.897774	S.D. dependent var	7.223449	
S.E. of regression	2.309539	Sum squared resid	149.3512	
Durbin-Watson stat	1.025955			
Unweighted Statistics including Random Effects				
R-squared	0.855189	Mean dependent var	15.04200	
Adjusted R-squared	0.850017	S.D. dependent var	7.223449	
S.E. of regression	2.797470	Sum squared resid	219.1235	
Durbin-Watson stat	0.699275			

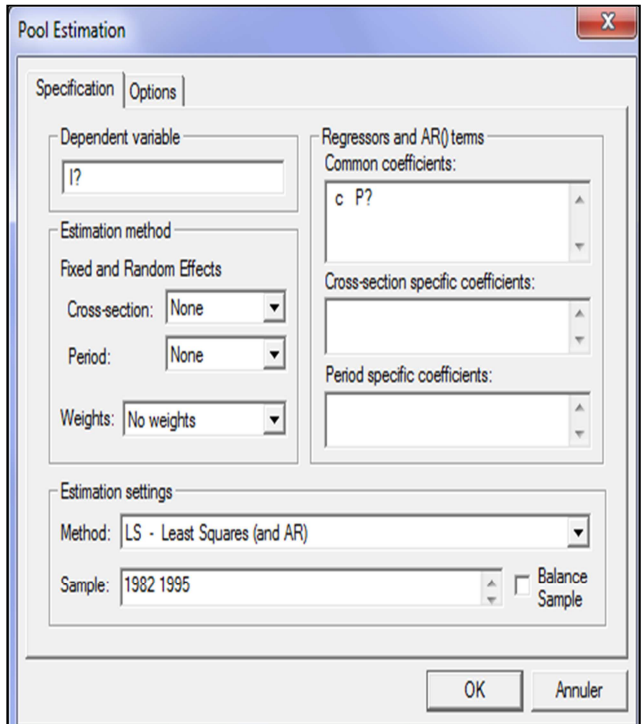


Dans l'output (Cfr Eviews 5), cliquer sur : View/Representations :

$I_FIRME1 = 0 - 0.7609358021 + 1.056157713 \cdot P_FIRME1$
 $I_FIRME2 = 0 - 0.7609358021 + 1.056157713 \cdot P_FIRME2$
 $I_FIRME3 = 0 - 0.7609358021 + 1.056157713 \cdot P_FIRME3$

► Estimation du modèle sans effets (Modèle 1)

Sur Eviews, Cfr la procédure pour estimer le modèle à effets fixes ou celui à effets aléatoires. Seulement : sélectionner « None » dans Cross section et dans Period. Aussi, dans common coefficient, ajouter la constante « c » en plus de la variable explicative « P » (Cfr figure à gauche ci-dessous) :



Dependent Variable: I?
 Method: Pooled Least Squares
 Date: 12/16/13 Time: 21:43
 Sample (adjusted): 1982 1991
 Included observations: 10 after adjustments
 Cross-sections included: 3
 Total pool (balanced) observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.760936	0.928620	-0.819426	0.4195
P?	1.056158	0.056970	18.53875	0.0000

R-squared	0.924667	Mean dependent var	15.04200
Adjusted R-squared	0.921977	S.D. dependent var	7.223449
S.E. of regression	2.017697	Akaike info criterion	4.306131
Sum squared resid	113.9908	Schwarz criterion	4.399544
Log likelihood	-62.59196	F-statistic	343.6853
Durbin-Watson stat	1.250819	Prob(F-statistic)	0.000000

Dans l'output, cliquer sur : View/Representations :

$I_FIRME1 = -0.7609358021 + 1.056157713 \cdot P_FIRME1$
 $I_FIRME2 = -0.7609358021 + 1.056157713 \cdot P_FIRME2$
 $I_FIRME3 = -0.7609358021 + 1.056157713 \cdot P_FIRME3$

Sur Stata, taper : `reg I P`

Source	SS	df	MS	Number of obs = 30		
Model	1399.17728	1	1399.17728	F(1, 28)	= 343.69	
Residual	113.990771	28	4.07109895	Prob > F	= 0.0000	
Total	1513.16805	29	52.1782087	R-squared	= 0.9247	
				Adj R-squared	= 0.9220	
				Root MSE	= 2.0177	
	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
P	1.056158	.0569703	18.54	0.000	.9394594	1.172856
_cons	-.7609357	.92862	-0.82	0.419	-2.663128	1.141256



b) Tests de spécification sur données de panel (test de Fisher/Likelihood Ratio Test ; test de Breusch et Pagan ; et test de Hausman)

► **Test de Fisher ou Likelihood Ratio Test : test de présence d'effets fixes (MCO vs Within)**

Soient les deux modèles :

$$\text{Modèle 1 : } I_{it} = c + b_1 P_{it} + e_{it} \dots \dots [5.2]$$

$$\text{Modèle 2 : } I_{it} = a_i + b_1 P_{it} + e_{it} \dots \dots [5.3]$$

Contrainte à tester : $a_1 = a_2 = a_3 = \dots = a_n = 0$.

Les hypothèses du test sont :

H_0 : Absence d'effets ($\text{prob} > 5\%$ ou $F_c < F_t$) : retenir le Modèle 1

H_1 : Présence d'effets fixes ($\text{prob} < 5\%$ ou $F_c > F_t$) : retenir le Modèle 2

Sur Stata, la statistique de Fisher est automatiquement calculée après estimation du modèle à effets fixes (commande : `xtreg I P, fe`), il suffit de lire la dernière ligne :

F test that all u_i=0:	F(2, 26) =	6.09	Prob > F = 0.0068
-------------------------------	-------------------	-------------	-----------------------------

Sur Eviews 6, après avoir estimé le modèle à effets fixes (dans l'output), cliquer sur : View → Fixed/Random effects Testing → **Redundant Fixed Effects-Likelihood Ratio** :

Redundant Fixed Effects Tests				
Pool: PANJ				
Test cross-section fixed effects				
Effects Test	Statistic	d.f.	Prob.	
Cross-section F	6.087574	(2,26)	0.0068	
Cross-section Chi-square	11.522646	2	0.0031	
Cross-section fixed effects test equation:				
Dependent Variable: I?				
Method: Panel Least Squares				
Date: 12/15/13 Time: 19:35				
Sample (adjusted): 1982 1991				
Included observations: 10 after adjustments				
Cross-sections included: 3				
Total pool (balanced) observations: 30				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.760936	0.928620	-0.819426	0.4195
P?	1.056158	0.056970	18.53875	0.0000
R-squared	0.924667	Mean dependent var	15.04200	
Adjusted R-squared	0.921977	S.D. dependent var	7.223449	
S.E. of regression	2.017697	Akaike info criterion	4.306131	
Sum squared resid	113.9908	Schwarz criterion	4.399544	
Log likelihood	-62.59196	Hannan-Quinn criter.	4.336014	
F-statistic	343.6853	Durbin-Watson stat	1.250819	
Prob(F-statistic)	0.000000			



_____ **Commentaire :** Parce que « $prob < 5\%$ », nous acceptons l'hypothèse alternative selon laquelle il y a présence d'effets fixes dans le modèle (retenir le Modèle 2).

► **Test de Breusch et Pagan (LM-test) : test de présence d'effets aléatoires (MCO vs MCG)**

Soient les deux modèles :

$$\text{Modèle 1 : } I_{it} = c + b_1 P_{it} + e_{it} \dots \dots [5.2]$$

$$\text{Modèle 3 : } I_{it} = \mu + b_1 P_{it} + a_i + e_{it} \dots \dots [5.4]$$

Les hypothèses du test sont :

H_0 : Absence d'effets ($prob > 5\%$ ou $F_c < F_t$) : retenir le Modèle 1

H_1 : Présence d'effets aléatoires ($prob < 5\%$ ou $F_c > F_t$) : retenir le Modèle 3

_____ Sur Stata, après estimation du modèle à effets aléatoires, faire : **xttest0**

Breusch and Pagan Lagrangian multiplier test for random effects:

I[firme,t] = Xb + u[firme] + e[firme,t]

Estimated results:

		Var	sd = sqrt(Var)
-----+-----			
I		52.17821	7.223449
e		2.985994	1.728003
u		0	0

Test: Var(u) = 0

chi2(1) = 6.81
Prob > chi2 = 0.0091

_____ Sur Eviews (3, 4, 5, 6 et 7), ce test n'est pas programmé. Nous concluons en faveur du modèle à effets aléatoires (accepter H_1 et retenir le Modèle 3).

► **Test de Hausman : test de présence d'effets aléatoires (MCG vs Within)**

Les hypothèses du test sont :

H_0 : Présence d'effets aléatoires ($prob > 5\%$ ou $H < \chi_k^2$)

H_1 : Présence d'effets fixes ($prob < 5\%$ ou $H > \chi_k^2$)

_____ Sur Stata, faire :

$\left\{ \begin{array}{l} \text{xtreg I P, re} \\ \text{est store re} \\ \text{xtreg I P, fe} \\ \text{hausman re} \end{array} \right.$

	---- Coefficients ----			
	(b) re	(B) .	(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
-----+-----				
P	1.056158	1.096293	-.0401357	.0269693
-----+-----				
b = consistent under H_0 and H_a ; obtained from xtreg				
B = inconsistent under H_a , efficient under H_0 ; obtained from xtreg				
Test: Ho: difference in coefficients not systematic				
	chi2(1) = (b-B)'[(V_b-V_B)^(-1)](b-B)			
	= 2.21			
	Prob>chi2 = 0.1367			



_____ **Sur EViews 6**, après avoir estimé le modèle à effets fixes (dans l'output), cliquer sur : View → Fixed/Random effects Testing → **Correlated Random Effects-Hausman Test** :

Correlated Random Effects - Hausman Test				
Pool: PANJ				
Test cross-section random effects				
Test Summary	Chi-Sq. Statistic	Chi-Sq. d.f.	Prob.	
Cross-section random	11.695490	1	0.0006	
** WARNING: estimated cross-section random effects variance is zero.				
Cross-section random effects test comparisons:				
Variable	Fixed	Random	Var(Diff.)	Prob.
P?	1.096293	1.056158	0.000138	0.0006
Cross-section random effects test equation:				
Dependent Variable: I?				
Method: Panel Least Squares				
Date: 12/15/13 Time: 19:53				
Sample (adjusted): 1982 1991				
Included observations: 10 after adjustments				
Cross-sections included: 3				
Total pool (balanced) observations: 30				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.361474	0.814448	-1.671651	0.1066
P?	1.096293	0.050182	21.84621	0.0000
Effects Specification				
Cross-section fixed (dummy variables)				
R-squared	0.948693	Mean dependent var	15.04200	
Adjusted R-squared	0.942773	S.D. dependent var	7.223449	
S.E. of regression	1.728003	Akaike info criterion	4.055376	
Sum squared resid	77.63586	Schwarz criterion	4.242202	
Log likelihood	-56.83064	Hannan-Quinn criter.	4.115143	
F-statistic	160.2517	Durbin-Watson stat	1.834669	
Prob(F-statistic)	0.000000			

_____ **Commentaire** : Les résultats du test d'Hausman diffèrent selon qu'on recourt à Stata 9 (le modèle retenu est celui à effets aléatoires, car : « $prob > 5\%$ ») ou à Eviews 6 (le modèle retenu est celui à effets fixes, car : « $prob < 5\%$ »). Nous retenons le modèle à effets fixes (il offre des paramètres plus significatifs).

c) Tests de stationnarité/racine unitaire sur données de panel et ceux de cointégration sur données de panel

- Tests de stationnarité/racine unitaire sur données de panel (test de Im-Pesaran-Shin (1997), celui de Levin-lin (1993), de Hadri, d'ADF, etc.)

Les hypothèses des tests sont :

H_0 : Présence de racine unitaire/Série non stationnaire ($prob > 5\%$)

H_1 : Absence de racine unitaire/Série stationnaire ($prob < 5\%$)

_____ Sur Stata (les commandes sont à télécharger : modules complémentaires).



Sur **Eviews 6**, dans l'output des données, cliquer sur : View/Unit Root Test... : (à gauche : la suite de la procédure et, à droite : les résultats du test). Les résultats suivants concernent « I » (avec trend and intercept) et « P » [avec trend and intercept—Cfr EViews 5.1 (le test de « Hadri » est intégré), et sans trend ni intercept sur EViews 6.1] : les variables « I et P » sont stationnaires.

Unit Root Test

Pool series: I?

Test type: Summary

Test for unit root in:

- ☒ Level
- ☐ 1st difference
- ☐ 2nd difference

Include in test equation:

- ☐ Individual intercept
- ☒ Individual intercept and trend
- ☐ None

Options:

- ☐ Use balanced sample

Lag length:

- ☒ Automatic selection: Schwarz
- ☐ User specified: 0

Maximum lags: *

(* - indicates automatic selection of maximum lags)

Spectral estimation:

Kernel method: Bartlett

Bandwidth selection:

- ☒ Automatic: Newey-West
- ☐ User specified: 2

OK Cancel

Pool unit root test: Summary

Series: I_FIRME1, I_FIRME2, I_FIRME3

Date: 12/15/13 Time: 20:45

Sample: 1982 1995

Exogenous variables: Individual effects, individual linear trends

Automatic selection of maximum lags

Automatic selection of lags based on SIC: 0 to 1

Newey-West bandwidth selection using Bartlett kernel

Method	Statistic	Prob.**	Cross-sections	Obs
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu t*	-5.13760	0.0000	3	25
Breitung t-stat	-0.92847	0.1766	3	22
Null: Unit root (assumes individual unit root process)				
Im, Pesaran and Shin W-stat	-1.05963	0.1447	3	25
ADF - Fisher Chi-square	14.7010	0.0227	3	25
PP - Fisher Chi-square	31.3318	0.0000	3	27

** Probabilities for Fisher tests are computed using an asymptotic Chi-square distribution. All other tests assume asymptotic normality.

Pool unit root test: Summary

Date: 12/25/13 Time: 12:35

Sample: 1982 1995

Series: P_FIRME1, P_FIRME2, P_FIRME3

Exogenous variables: Individual effects, individual linear trends

Automatic selection of maximum lags

Automatic selection of lags based on SIC: 0 to 1

Newey-West bandwidth selection using Bartlett kernel

Method	Statistic	Prob.**	Cross-sections	Obs
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu t*	-5.54090	0.0000	3	37
Breitung t-stat	-1.75913	0.0393	3	34
Null: Unit root (assumes individual unit root process)				
Im, Pesaran and Shin W-stat	-3.82076	0.0001	3	37
ADF - Fisher Chi-square	24.1514	0.0005	3	37
PP - Fisher Chi-square	42.5277	0.0000	3	39

Null: No unit root (assumes common unit root process)

Hadri Z-stat 3.54004 0.0002 3 42

** Probabilities for Fisher tests are computed using an asymptotic Chi-square distribution. All other tests assume asymptotic normality.

Pool unit root test: Summary

Series: P_FIRME1, P_FIRME2, P_FIRME3

Date: 12/15/13 Time: 20:58

Sample: 1982 1995

Exogenous variables: None

Automatic selection of maximum lags

Automatic selection of lags based on SIC: 0 to 2

Newey-West bandwidth selection using Bartlett kernel

Method	Statistic	Prob.**	Cross-sections	Obs
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu t*	-0.08291	0.4670	3	35
Null: Unit root (assumes individual unit root process)				
ADF - Fisher Chi-square	4.22284	0.6465	3	35
PP - Fisher Chi-square	8.03929	0.2352	3	39

** Probabilities for Fisher tests are computed using an asymptotic Chi-square distribution. All other tests assume asymptotic normality.



► Tests de cointégration sur données de panel (test de Pedroni, celui de Kao et celui de Fisher)

Les hypothèses des tests sont :

H_0 : Les séries ne sont pas cointégrées ($prob > 5\%$)

H_1 : Les séries sont cointégrées ($prob < 5\%$)

_____ Sur Stata (les commandes sont à télécharger : modules complémentaires).

_____ Sur **Eviews 6**, dans l'output des données, cliquer sur : View/Cointegration Test... : (à gauche : la suite de la procédure et, à droite : les résultats des tests).

Pedroni Residual Cointegration Test					
Series: I? P?					
Date: 12/16/13 Time: 22:46					
Sample: 1982 1995					
Included observations: 14					
Cross-sections included: 3					
Null Hypothesis: No cointegration					
Trend assumption: Deterministic intercept and trend					
Lag selection: fixed at 1					
Newey-West bandwidth selection with Bartlett kernel					
Alternative hypothesis: common AR coeffs. (within-dimension)					
	Statistic	Prob.	Weighted Statistic	Prob.	
Panel v-Statistic	-1.047573	0.8526	-1.544106	0.9387	
Panel rho-Statistic	0.516791	0.6973	0.278993	0.6099	
Panel PP-Statistic	-1.105975	0.1344	-3.239651	0.0006	
Panel ADF-Statistic	-0.197424	0.4217	-1.033363	0.1507	
Alternative hypothesis: individual AR coeffs. (between-dimension)					
	Statistic	Prob.			
Group rho-Statistic	1.166881	0.8784			
Group PP-Statistic	-2.927485	0.0017			
Group ADF-Statistic	-0.261599	0.3968			
Cross section specific results					
Phillips-Perron results (non-parametric)					
Cross ID	AR(1)	Variance	HAC	Bandwidth	Obs
_FIRME1	-0.370	1.029504	0.299911	8.00	9
_FIRME2	0.263	1.725147	1.725147	0.00	9
_FIRME3	0.152	3.722207	3.769092	1.00	9
Augmented Dickey-Fuller results (parametric)					
Cross ID	AR(1)	Variance	Lag	Max lag	Obs
_FIRME1	-0.811	0.955338	1	--	8
_FIRME2	0.078	1.442996	1	--	8
_FIRME3	0.002	3.955053	1	--	8

Note: Selon les tests « Rho », « V » et « ADF », les variables « I et P » ne sont pas cointégrées ($prob > 5\%$), alors que le test de Philippe et Perron (PP) accepte l'hypothèse de cointégration ($prob < 5\%$) pour « I et P ».

Ci-dessous, le résultat du tests de Kao (Engle-Granger based) et celui du test de Fisher (Combined Johansen) → (sélectionner ces tests dans « Test type » : Cfr figure à gauche ci-haut) :



Kao Residual Cointegration Test				
Series: I? P?				
Date: 12/16/13 Time: 23:03				
Sample: 1982 1995				
Included observations: 14				
Null Hypothesis: No cointegration				
Trend assumption: No deterministic trend				
Lag selection: fixed at 1				
Newey-West bandwidth selection using Bartlett kernel				
ADF		t-Statistic	Prob.	
		-2.053427	0.0200	
Residual variance		4.878552		
HAC variance		1.223970		
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(RESID?)				
Method: Panel Least Squares				
Date: 12/16/13 Time: 23:03				
Sample (adjusted): 1984 1991				
Included observations: 8 after adjustments				
Cross-sections included: 3				
Total pool (balanced) observations: 24				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
RESID?(-1)	-1.032520	0.307011	-3.363137	0.0028
D(RESID?(-1))	0.130770	0.217271	0.601877	0.5534
R-squared	0.434755	Mean dependent var	0.164603	
Adjusted R-squared	0.409062	S.D. dependent var	2.322191	
S.E. of regression	1.785127	Akaike info criterion	4.076512	
Sum squared resid	70.10693	Schwarz criterion	4.174683	
Log likelihood	-46.91814	Hannan-Quinn criter.	4.102556	
Durbin-Watson stat	2.053960			

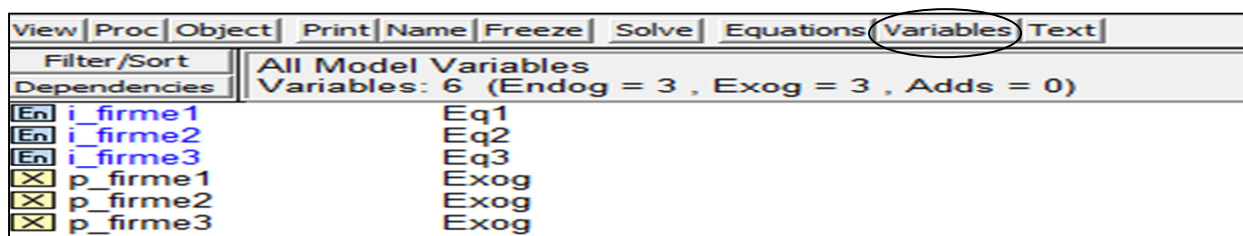
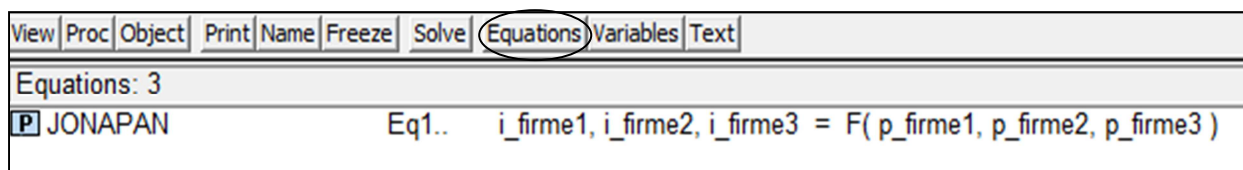
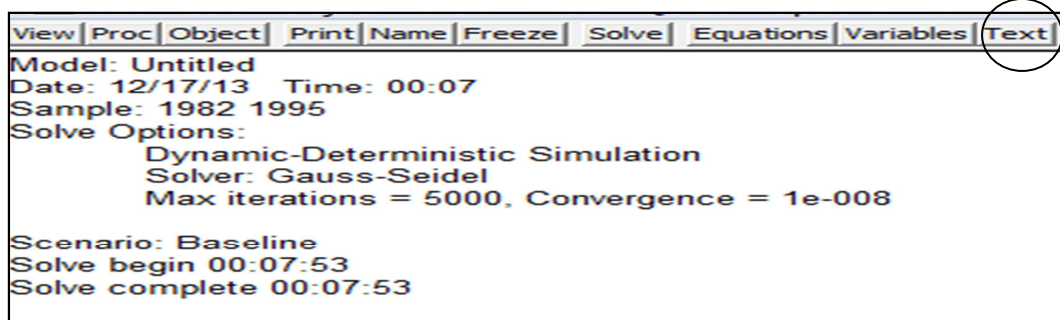
Johansen Fisher Panel Cointegration Test				
Series: I? P?				
Date: 12/16/13 Time: 23:07				
Sample: 1982 1995				
Included observations: 14				
Trend assumption: Quadratic deterministic trend				
Lags interval (in first differences): 1 1				
Unrestricted Cointegration Rank Test (Trace and Maximum Eigenvalue)				
Hypothesized	Fisher Stat.*	Prob.	Fisher Stat.*	Prob.
No. of CE(s)	(from trace test)		(from max-eigen test)	
None	40.43	0.0000	36.01	0.0000
At most 1	16.39	0.0118	16.39	0.0118
* Probabilities are computed using asymptotic Chi-square distribution.				
Individual cross section results				
Cross Section	Trace Test	Prob.**	Max-Eign Test	Prob.**
	Statistics		Statistics	
Hypothesis of no cointegration				
_FIRME1	28.1105	0.0016	23.3727	0.0055
_FIRME2	27.1105	0.0024	25.6634	0.0023
_FIRME3	31.4697	0.0004	27.2910	0.0012
Hypothesis of at most 1 cointegration relationship				
_FIRME1	4.7378	0.0295	4.7378	0.0295
_FIRME2	1.4471	0.2290	1.4471	0.2290
_FIRME3	4.1787	0.0409	4.1787	0.0409
**Mackinnon-Haug-Michelis (1999) p-values				

d) *Prévision (à horizon de 4 ans) des modèles estimés (Modèle 2 et Modèle 3)*

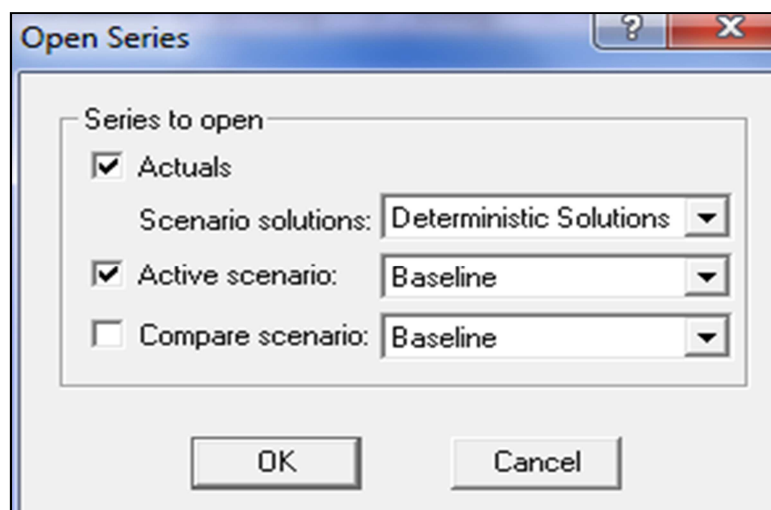
Sur EViews 5.1, dans l'output des résultats (relatif au modèle à effets fixes estimé), cliquer : **Proc/Make Model/Solve** : la boîte de dialogue ci-dessous apparaît (elle complète la procédure) :



_____ Cliquer sur « OK » pour obtenir les résultats (affichages) ci-dessous selon les onglets choisis :

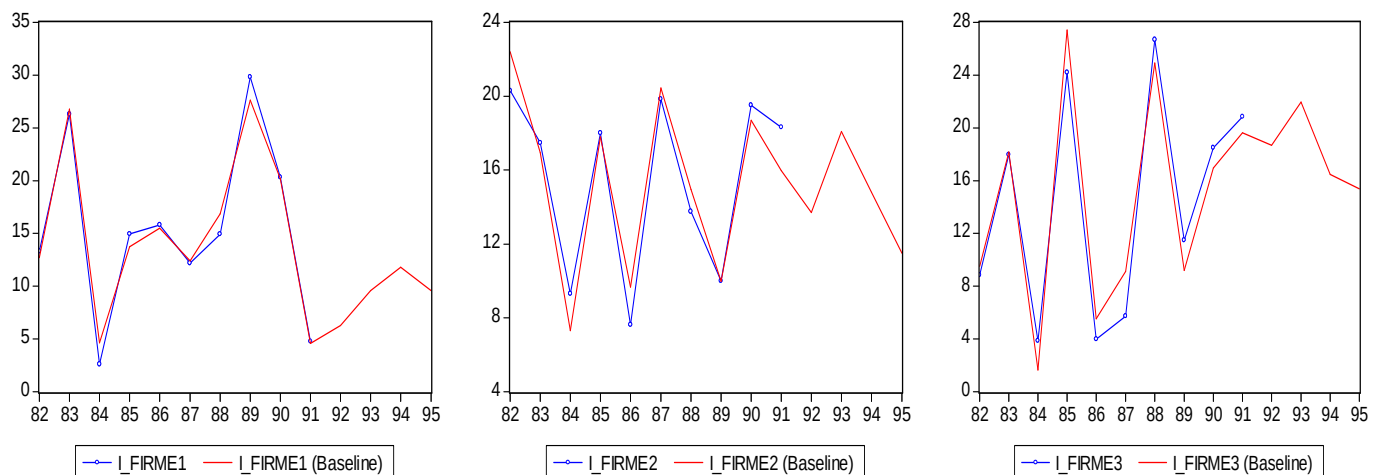


_____ En double cliquant sur une rubrique (ex: *i_firme1*), la boîte de dialogue suivante s'affiche (la figure complète la procédure) :



_____ Cliquer sur « OK » pour afficher les données réelles et prévues de la rubrique (variable) choisie. Pour obtenir les graphiques ci-dessous, dans l'output des données affiché, cliquer sur : **View/Graph/Line** :





Sur **EViews 3.1**, dans l'output des résultats (relatif au modèle à effets aléatoires estimé), Cliquer : **Procs/Make Model** (la figure ci-dessous apparaît) :

```

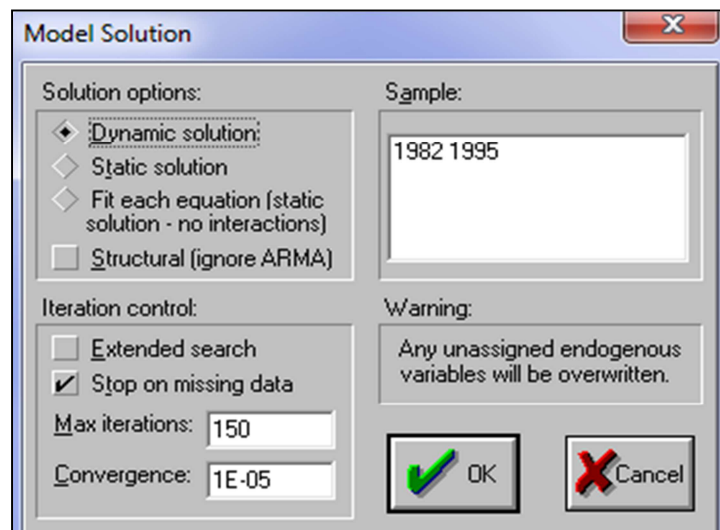
EViews - [Model: UNTITLED Workfile: UNTITLED]
File Edit Objects View Procs Quick Options Window Help
View Procs Objects Print Name Freeze Solve Merge Spec Endog

ASSIGN @ALL F

I_FIRME1 = -0.01969018251 - 0.1795886236 + 1.017304533*P_FIRME1
I_FIRME2 = 1.357056859 - 0.1795886236 + 1.017304533*P_FIRME2
I_FIRME3 = -1.337366677 - 0.1795886236 + 1.017304533*P_FIRME3

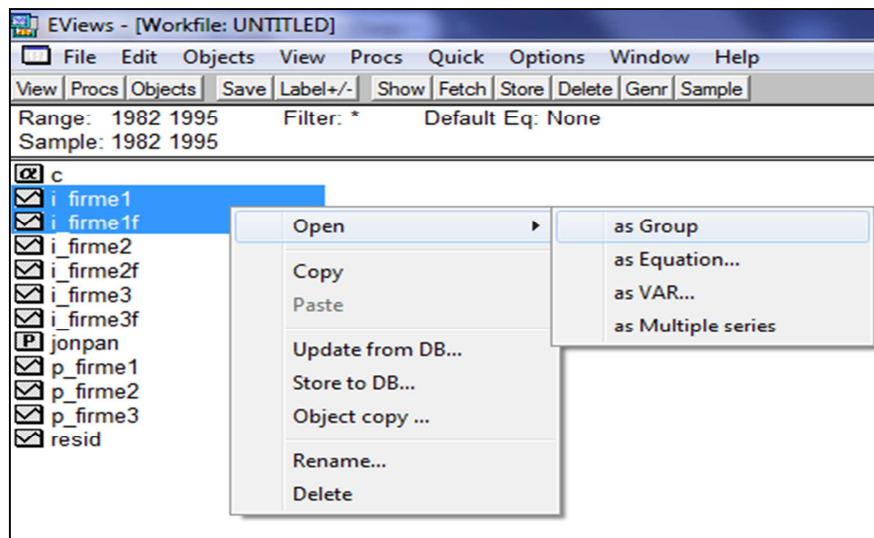
```

Ensuite, cliquer sur l'onglet **Solve** (Cfr figure précédente) pour afficher la boîte de dialogue qui suit (elle complète la procédure) :

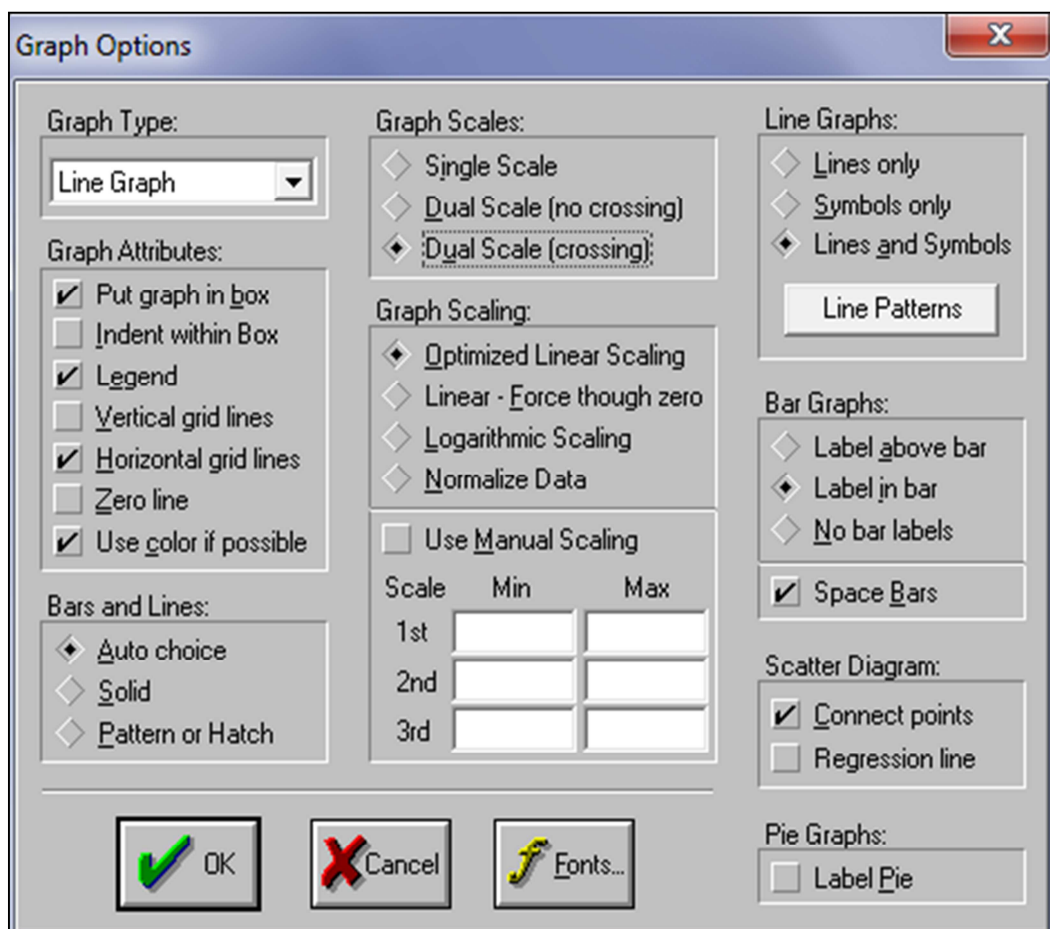


Cliquer sur « **OK** » pour produire les valeurs prédites (variables avec l'indice « f ») à côté des valeurs réelles. Dans l'espace de travail, sélectionner deux variables (réelles et prédites. Ex : *i_firme1* et *i_firme2*) → **clic droit** → **Open** → **as Group**.



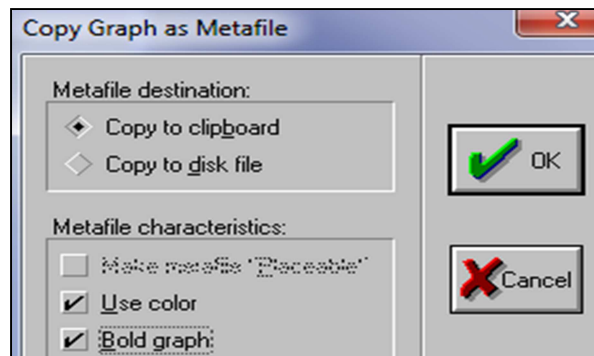


Dans l'output des données ainsi obtenu, suivre : **View/Graph/Line**, pour représenter graphiquement les variables « *i_firme1* et *i_firme2* ». Dans l'output du graphique, cliquer : **Objects/View Options/Options...**, pour obtenir la boîte de dialogue qui suit (elle permet de modifier le graphique dans sa présentation/affichage) :

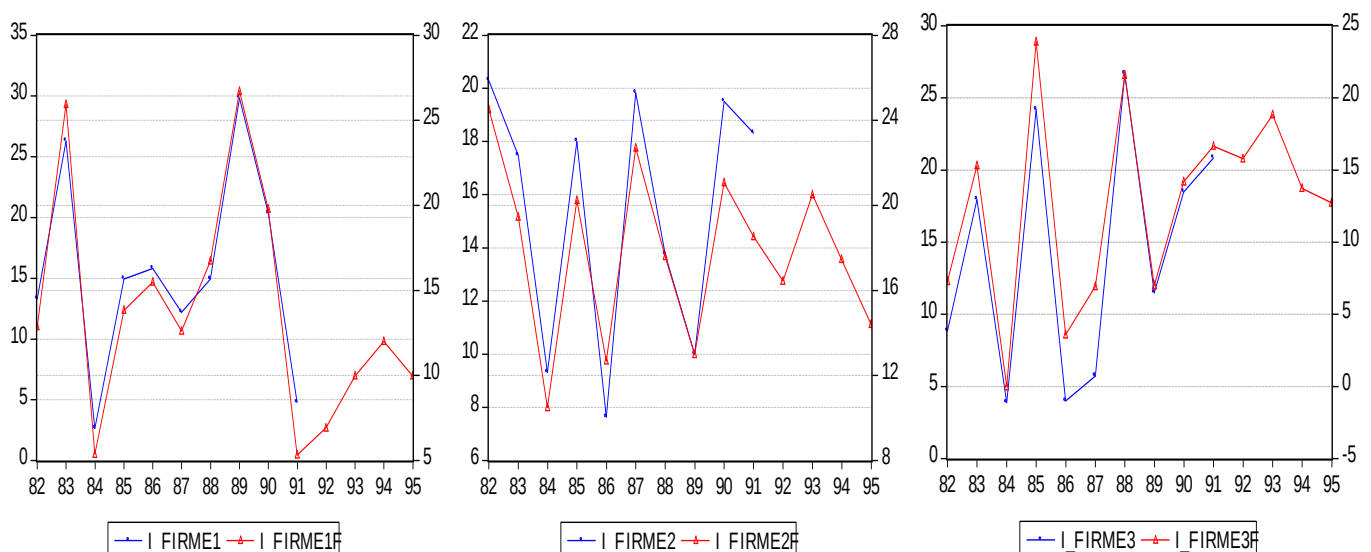


Après avoir cliqué sur « **OK** », faites « **CTRL+C** », ce qui ouvre la boîte de dialogue ci-dessous (elle complète la procédure) :





Cliquer sur « OK », ensuite « CTRL+V » dans Word, ce qui amène aux graphiques ci-dessous (reprendre la procédure pour d'autres variables si l'on tient à reproduire les deux autres graphiques) :



Les valeurs réelles et prédites (prévues pour les années 1992 à 1995) à l'aide du **modèle à effets aléatoires (sur EViews 3.1)** se présentent comme suit :

obs	I_FIRME1	I_FIRME1F	I_FIRME2	I_FIRME2F	I_FIRME3	I_FIRME3F
1982	13.32000	12.87308	20.30000	24.50426	8.850000	7.282729
1983	26.30000	25.93527	17.47000	19.44826	17.96000	15.31943
1984	2.620000	5.375550	9.310000	10.49598	3.870000	-0.021518
1985	14.94000	13.82935	18.01000	20.23158	24.19000	23.88514
1986	15.80000	15.47738	7.630000	12.68318	3.990000	3.579740
1987	12.20000	12.60859	19.84000	22.69346	5.730000	6.926672
1988	14.93000	16.72867	13.76000	17.58659	26.68000	21.57586
1989	29.82000	26.70843	10.00000	12.98837	11.49000	6.987711
1990	20.32000	19.78058	19.51000	21.06577	18.49000	14.19023
1991	4.770000	5.324685	18.32000	18.53268	20.84000	16.66228
1992	NA	6.921853	NA	16.43704	NA	15.77722
1993	NA	9.973767	NA	20.50625	NA	18.82914
1994	NA	12.00838	NA	17.45434	NA	13.74261
1995	NA	9.973767	NA	14.40243	NA	12.72531

Les valeurs réelles et prédites (prévues pour les années 1992 à 1995) à l'aide du **modèle à effets fixes (Sur EViews 5.1)** se présentent comme suit :



Obs	I_FIRME1	I_FIRME1_0	I_FIRME2	I_FIRME2_0	I_FIRME3	I_FIRME3_0
1982	13.32000	12.70974	20.30000	22.39949	8.850000	9.514671
1983	26.30000	26.78615	17.47000	16.95091	17.96000	18.17539
1984	2.620000	4.630058	9.310000	7.303525	3.870000	1.643284
1985	14.94000	13.74026	18.01000	17.79505	24.19000	27.40618
1986	15.80000	15.51625	7.630000	9.660556	3.990000	5.524163
1987	12.20000	12.42470	19.84000	20.44808	5.730000	9.130969
1988	14.93000	16.86469	13.76000	14.94469	26.68000	24.91759
1989	29.82000	27.61933	10.00000	9.989444	11.49000	9.196746
1990	20.32000	20.15357	19.51000	18.69401	18.49000	16.95850
1991	4.770000	4.575243	18.32000	15.96424	20.84000	19.62250
1992	NA	6.296424	NA	13.70588	NA	18.66872
1993	NA	9.585304	NA	18.09105	NA	21.95760
1994	NA	11.77789	NA	14.80217	NA	16.47613
1995	NA	9.585304	NA	11.51329	NA	15.37984

Sur Stata, faire :

$\left\{ \begin{array}{l} \text{xtreg I P, fe} \\ \text{predict IF, xb} \\ \text{xtreg I P, re} \\ \text{predict IA, xb} \end{array} \right.$

- Avec « IF » : Investissement prévu à l'aide du modèle à effets fixes ;
- Avec « IA » : Investissement prévu à l'aide du modèle à effets aléatoires.

Années	Firmes	IF	IA
1992	1	6.312581	6.632168
1993	1	9.601461	9.800641
1994	1	11.79405	11.91296
1995	1	9.601461	9.800641
1992	2	15.08293	15.08143
1993	2	19.4681	19.30606
1994	2	16.17922	16.13759
1995	2	12.89034	12.96911
1992	3	17.27551	17.19374
1993	3	20.5644	20.36222
1994	3	15.08293	15.08143
1995	3	13.98664	14.02527

e) Estimation d'un modèle à correction d'erreur à effets fixes

Sous l'hypothèse que les variables « I » et « P » sont cointégrées (intégrées de même ordre et non stationnaire de même type/DS), nous nous proposons d'estimer un modèle à correction d'erreur à effets fixes à la manière de HANDRY, spécifié comme suit :

$$DLI_{it} = a_i + b_1 DLP_{it} + b_2 LI_{it-1} + b_3 LP_{it-1} + e_{it} \dots \dots [5.3]'$$

Avec :

- o DLI_{it} et DLP_{it} : les variables « lit » et « Pit » prises en différences premières ;
- o LI_{it} et LP_{it} : les variables « lit » et « Pit » prises en logarithme ;
- o b_1 : l'élasticité à court terme ;



- o b_2 : la force de rappel ou le coefficient de correction d'erreur indiquant la vitesse avec laquelle tout déséquilibre entre les niveaux désiré et réel de « lit » est résorbé dans l'année qui suit le choc ;
- o $\frac{-b_3}{b_2}$: élasticité de long terme.

Sur **EViews 5.1**, dans l'output des données, cliquer sur l'onglet **PoolGenr** et taper les expressions ci-dessous l'une après l'autre en cliquant chaque fois sur « **OK** » et en allant de nouveau sur « **PoolGenr** » (pour générer les variables en différences premières et en logarithme) :

$\left\{ \begin{array}{l} \text{LI?}=\log(\text{I?}) \\ \text{DLI?}=\text{d}(\text{LI?}) \\ \text{LP?}=\log(\text{P?}) \\ \text{DLP?}=\text{d}(\text{LP?}) \end{array} \right.$

Ensuite, cliquer sur l'onglet **Estimate** et taper :

- o Dans « *Dependant variable* » : **DLI?**
- o Dans « *Regressors and AR() terms* » : **DLP? LI?(-1) LP?(-1)**

NB : Ne pas oublier de sélectionner le modèle à effets fixes dans “Estimation method”. Cliquer sur « **OK** » pour obtenir l'output ci-dessous :

Dependent Variable: DLI?				
Method: Pooled Least Squares				
Date: 12/22/13 Time: 13:51				
Sample (adjusted): 1983 1991				
Included observations: 9 after adjustments				
Cross-sections included: 3				
Total pool (balanced) observations: 27				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.015163	0.483370	-0.031370	0.9753
DLP?	0.972912	0.106000	9.178455	0.0000
LI?(-1)	-1.018984	0.219508	-4.642124	0.0001
LP?(-1)	1.021162	0.287093	3.556904	0.0019
Fixed Effects (Cross)				
_FIRME1--C	-0.075073			
_FIRME2--C	-0.076739			
_FIRME3--C	0.151812			
Effects Specification				
Cross-section fixed (dummy variables)				
R-squared	0.934564	Mean dependent var	-0.010114	
Adjusted R-squared	0.918985	S.D. dependent var	1.033086	
S.E. of regression	0.294049	Akaike info criterion	0.582991	
Sum squared resid	1.815764	Schwarz criterion	0.870955	
Log likelihood	-1.870380	F-statistic	59.98534	
Durbin-Watson stat	2.178162	Prob(F-statistic)	0.000000	

Commentaire : Le modèle à correction d'erreur à effets fixes estimé semble non valide, car la force de rappel « b_2 » estimée (coefficient associé à LI?(-1)) est statistiquement significatif, mais n'est pas compris entre 0 et 1.

Contournons le problème en prenant cette fois-ci les variables en niveau (exclure le logarithme). Pour ce faire (**Cfr EViews 3.1**), dans « **PoolGenr** », taper cette fois :



$$\begin{cases} DI? = d(I?) \\ DP? = d(P?) \end{cases}$$

Ensuite, cliquer sur l'onglet **Estimate** et taper :

- o Dans « *Dependant variable* » : **DI?**
- o Dans « *Regressors and AR() terms* » : **DP? I?(-1) P?(-1)**

NB : Ne pas oublier de sélectionner le modèle à effets fixes dans “ *Intercept* ”. Cliquer sur « **OK** » pour obtenir l'output ci-dessous (à gauche pour EViews 5.1 et à droite pour EViews 3.1 : la constante est obligatoire sur EViews 5.1) :

Dependent Variable: DI?				
Method: Pooled Least Squares				
Date: 12/22/13 Time: 14:30				
Sample (adjusted): 1983 1991				
Included observations: 9 after adjustments				
Cross-sections included: 3				
Total pool (balanced) observations: 27				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.711660	1.691985	-1.602650	0.1239
DP?	1.135481	0.061322	18.51680	0.0000
I?(-1)	-0.927142	0.216282	-4.286726	0.0003
P?(-1)	1.118548	0.265833	4.207703	0.0004
Fixed Effects (Cross)				
_FIRME1--C	-0.285082			
_FIRME2--C	-1.234882			
_FIRME3--C	1.519963			
Effects Specification				
Cross-section fixed (dummy variables)				
R-squared	0.982265	Mean dependent var	0.054074	
Adjusted R-squared	0.978043	S.D. dependent var	12.14586	
S.E. of regression	1.799763	Akaike info criterion	4.206317	
Sum squared resid	68.02210	Schwarz criterion	4.494281	
Log likelihood	-50.78528	F-statistic	232.6258	
Durbin-Watson stat	2.074516	Prob(F-statistic)	0.000000	

Dependent Variable: DI?				
Method: Pooled Least Squares				
Date: 12/22/13 Time: 14:30				
Sample(adjusted): 1983 1991				
Included observations: 9 after adjusting endpoints				
Number of cross-sections used: 3				
Total panel (balanced) observations: 27				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DP?	1.135481	0.061322	18.51680	0.0000
I?(-1)	-0.927142	0.216282	-4.286726	0.0003
P?(-1)	1.118548	0.265833	4.207703	0.0004
Fixed Effects				
_FIRME1--C	-2.996742			
_FIRME2--C	-3.946542			
_FIRME3--C	-1.191697			
R-squared	0.982265	Mean dependent var	0.054074	
Adjusted R-squared	0.978043	S.D. dependent var	12.14586	
S.E. of regression	1.799763	Sum squared resid	68.02210	
Log likelihood	-50.78528	F-statistic	581.5644	
Durbin-Watson stat	2.074516	Prob(F-statistic)	0.000000	

Commentaire : Le modèle à correction d'erreur à effets fixes estimé est valide, car la force de rappel « b_2 » estimée (coefficient associé à $I?(-1)$) est statistiquement significatif et est compris entre 0 et 1 (soit $\hat{b}_2 = -0,927$).

Sur Stata, faire :

$$\begin{cases} \text{gen } DI = d.I \\ \text{gen } DP = d.P \\ \text{xtreg } DI \text{ } DP \text{ } L.I \text{ } L.P, \text{ fe} \end{cases}$$



Fixed-effects (within) regression			Number of obs = 27			
Group variable (i): firme			Number of groups = 3			
R-sq: within = 0.9822			Obs per group: min = 9			
between = 0.0055			avg = 9.0			
overall = 0.9731			max = 9			
			F(3,21) = 385.19			
corr(u_i, Xb) = -0.0334			Prob > F = 0.0000			

DI	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	

DP	1.135481	.0613217	18.52	0.000	1.007956	1.263006
I						
L1.	-.9271422	.2162822	-4.29	0.000	-1.376926	-.4773589
P						
L1.	1.118548	.2658333	4.21	0.000	.5657169	1.671378
_cons	-2.711659	1.691985	-1.60	0.124	-6.230335	.8070164

sigma_u	1.3993731					
sigma_e	1.7997632					
rho	.3767745	(fraction of variance due to u_i)				

F test that all u_i=0:			F(2, 21) =	3.10	Prob > F = 0.0659	

f) Estimation d'un panel dynamique

_____ Sur Eviews : pas de programmes/commandes appropriés ;

_____ Sur Stata, la commande et les résultats se présentent comme suit :

Commande : xtabond DI DP L.I L.P, robust small

Arellano-Bond dynamic panel-data estimation			Number of obs	=	21
Group variable (i): firme			Number of groups	=	3
			F(2, 16)	=	16.02
Time variable (t): annee			Obs per group: min	=	7
			avg	=	7
			max	=	7
One-step results					

D.DI	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]

DI					
LD.	.0386632	.0678957	0.57	0.577	-.1052692 .1825956
DP					
D1.	1.08303	.0600337	18.04	0.000	.9557638 1.210295
I					
LD.	-.9719573	.1717465	-5.66	0.000	-1.336044 -.6078711
P					
LD.	1.009266	.2365404	4.27	0.001	.5078227 1.510709
_cons	.2461897	.1394549	1.77	0.097	-.0494415 .5418208

Arellano-Bond test that average autocovariance in residuals of order 1 is 0:					
H0: no autocorrelation z = -1.56 Pr > z = 0.1198					
Arellano-Bond test that average autocovariance in residuals of order 2 is 0:					
H0: no autocorrelation z = 1.31 Pr > z = 0.1916					

L'estimation de notre modèle à correction d'erreurs à effets fixes (modèle [5.3]') par l'approche d'Arellano et Bond (1991) --- recours à la méthode GMM/méthode de moments généralisés --- améliore les résultats en ce que les « t de student » produits sont robustes (corrigés d'hétéroscédasticité ou d'éventuelles autocorrélations des erreurs : à titre illustratif, la constante devient significative à 10%).



g) Tests post-estimation sur panel (inférence Statistique)

NB : Pour de raisons d'espace, nos tests ne se rapportent qu'au modèle à effets fixes estimé. Autant dire qu'il faut d'abord estimer le modèle à effets fixes avant d'effectuer tel ou tel autre test ci-dessous :

► **Tests de normalité des résidus :**

_____ Sur Eviews (les commandes ne sont pas programmées)

_____ Sur Stata, faire (test de Shapiro-Wilk) :

```
{ predict res, e
  swilk res
```

Shapiro-Wilk W test for normal data					
Variable	Obs	W	V	z	Prob>z
res	30	0.95530	1.421	0.726	0.23381

Autrement (Test de Jarque-Berra) :

Commande : sktest res

Skewness/Kurtosis tests for Normality				
Variable	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	joint Prob>chi2
res	0.374	0.436	1.50	0.4713

► **Test d'hétéroscédasticité de White :** il y a absence d'hétéroscédasticité (prob>5%).

_____ Sur Stata, faire :

```
{ gen res2=res^2
  reg res2 P
```

Source	SS	df	MS	Number of obs = 30		
Model	.827210215	1	.827210215	F(1, 28) =	0.09	
Residual	256.493684	28	9.1604887	Prob > F =	0.7660	
Total	257.320894	29	8.87313427	R-squared =	0.0032	
				Adj R-squared =	-0.0324	
				Root MSE =	3.0266	
res2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
P	-.0256803	.0854578	-0.30	0.766	-.2007327	.1493721
_cons	2.972108	1.392969	2.13	0.042	.1187394	5.825476

_____ Sur EViews, pas de commandes y relatives.



h) Estimation d'un modèle de panel à effets between et d'un modèle à effets aléatoires avec l'estimateur ml (maximum de vraisemblance)

Modèle à effets between

Commande : xtreg I P, be

```
Between regression (regression on group means)   Number of obs       =       30
Group variable (i): annee                       Number of groups    =       10
R-sq:  within = 0.8997                          Obs per group: min =        3
        between = 0.9577                        avg =       3.0
        overall = 0.9247                        max =        3
                                                F(1,8)              =    181.02
sd(u_i + avg(e_i.))= 1.077036                  Prob > F             =    0.0000
```

I	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
P	1.075904	.0799664	13.45	0.000	.8915016	1.260307
_cons	-1.056399	1.244041	-0.85	0.420	-3.925162	1.812364

Modèle à effets aléatoires avec l'estimateur ML

Commande : xtreg I P, mle

```
Fitting constant-only model:
Iteration 0:  log likelihood = -234.46955
Iteration 1:  log likelihood = -164.32753
Iteration 2:  log likelihood = -128.25246
Iteration 3:  log likelihood = -111.04199
Iteration 4:  log likelihood = -103.94225
Iteration 5:  log likelihood = -101.75884
Iteration 6:  log likelihood = -101.39617
Iteration 7:  log likelihood = -101.37966
Iteration 8:  log likelihood = -101.37961
```

```
Fitting full model:
Iteration 0:  log likelihood = -62.626851
Iteration 1:  log likelihood = -62.592274
Iteration 2:  log likelihood = -62.591958
```

```
Random-effects ML regression               Number of obs       =       30
Group variable (i): annee                 Number of groups    =       10
Random effects u_i ~ Gaussian              Obs per group: min =        3
                                                avg =       3.0
                                                max =        3
                                                LR chi2(1)         =    77.58
Log likelihood = -62.591958                Prob > chi2         =    0.0000
```

I	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
P	1.056158	.0550371	19.19	0.000	.948287	1.164028
_cons	-.7609357	.8971088	-0.85	0.396	-2.519237	.9973652
/sigma_u	0	.695689			.	.
/sigma_e	1.94928	.2516345			1.513532	2.51048
rho	0	.			.	.

```
Likelihood-ratio test of sigma_u=0: chibar2(01)= 0.00 Prob>=chibar2 = 1.000
```



i) Estimation d'un panel corrigé d'hypothèses violées

NB : Pour de raisons d'espace, nous ne nous basons que sur le modèle à effets fixes. Aussi, seul Stata est utilisé (EViews n'offre pas des méthodes de correction d'hypothèses violées sur modèle de panel).

Sur Stata, suivre les commandes appropriées :

- Correction de l'hétéroscédasticité sur un modèle à effets fixes (commande Stata: areg)

Commande : areg I P, absorb(firme) robust

Linear regression, absorbing indicators					Number of obs = 30		
					F(1, 26) = 320.32		
					Prob > F = 0.0000		
					R-squared = 0.9487		
					Adj R-squared = 0.9428		
					Root MSE = 1.728		

I	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]		

P	1.096293	.0612536	17.90	0.000	.9703849	1.222202	
_cons	-1.361473	.9627057	-1.41	0.169	-3.340343	.6173967	

firmes	absorbed					(3 categories)	

- Correction de l'autocorrélation des erreurs d'ordre 1 sur un modèle à effets fixes (commande Stata: xtregar)

Commande : xtregar I P, fe

FE (within) regression with AR(1) disturbances				Number of obs	=	20
Group variable (i): annee				Number of groups	=	10
R-sq:	within	=	0.8349	Obs per group:	min	= 2
	between	=	0.9274		avg	= 2.0
	overall	=	0.8935		max	= 2
				F(1,9)	=	45.50
corr(u_i, Xb) = 0.2875				Prob > F	=	0.0001

I	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	

P	.9204059	.1364445	6.75	0.000	.6117469	1.229065
_cons	1.245292	2.056354	0.61	0.560	-3.406503	5.897086

rho_ar	.01041954					
sigma_u	1.6312589					
sigma_e	2.330513					
rho_fov	.32883194	(fraction of variance due to u_i)				

F test that all u_i=0:		F(9,9)	=	0.89	Prob > F = 0.5646	



Sur EViews 5.1, dans l'output des résultats, cliquer sur l'onglet **Estimate** → dans « Weights », cocher « Cross section SUR » pour corriger le modèle à effets fixes estimé d'éventuelles autocorrélations d'erreurs et/ou d'hétéroscédasticité, grâce à la méthode SUR/Seemingly Unrelated Estimator (à gauche : la figure qui illustre et complète la procédure ; et à droite : les résultats) :

Dependent Variable: I?
Method: Pooled EGLS (Cross-section SUR)
Date: 12/25/13 Time: 14:16
Sample (adjusted): 1982 1991
Included observations: 10 after adjustments
Cross-sections included: 3
Total pool (balanced) observations: 30
Linear estimation after one-step weighting matrix

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.764662	0.675264	-2.613293	0.0147
P?	1.123240	0.041806	26.86779	0.0000
Fixed Effects (Cross)				
_FIRME1--C	-0.027860			
_FIRME2--C	-1.420065			
_FIRME3--C	1.447926			
Effects Specification				
Cross-section fixed (dummy variables)				
Weighted Statistics				
R-squared	0.972542	Mean dependent var	10.47058	
Adjusted R-squared	0.969374	S.D. dependent var	6.095359	
S.E. of regression	1.066709	Sum squared resid	29.58455	
F-statistic	306.9672	Durbin-Watson stat	1.995330	
Prob(F-statistic)	0.000000			
Unweighted Statistics				
R-squared	0.948124	Mean dependent var	15.04200	
Sum squared resid	78.49683	Durbin-Watson stat	1.892121	

Commentaire : Le modèle à effets fixes corrigé par la méthode SUR paraît plus significatif que tout autre modèle (la constante devient significative à 2%).

j) Annexes : Statistiques descriptives, évolutions graphiques des variables et évolutions graphiques des résidus

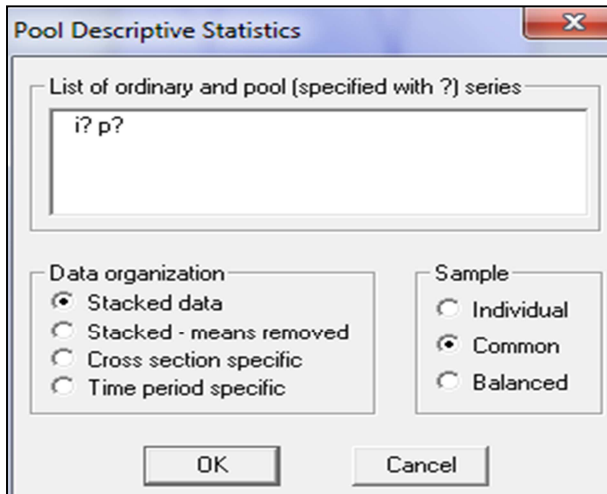
► Statistiques descriptives de données en panel

Sur Stata, faire : **xtsum I P**

Variable		Mean	Std. Dev.	Min	Max	Observations
I	overall	15.042	7.223449	2.62	29.82	N = 30
	between		4.93589	5.266667	20.57667	n = 10
	within		5.430967	5.168667	27.75867	T = 3
P	overall	14.68762	5.893947	1.47	26.45	N = 42
	between		3.856906	5.37	20.06667	n = 14
	within		4.537459	6.664285	25.66429	T = 3



Sur EViews 5.1, dans l'output des résultats/données, suivre : View/Descriptive Statistics... → la boîte de dialogue suivante apparaît (elle complète la procédure) → cliquer sur « OK » pour obtenir les résultats à droite :



	I?	P?
Mean	15.04200	14.96267
Median	15.37000	15.78500
Maximum	29.82000	26.45000
Minimum	2.620000	1.470000
Std. Dev.	7.223449	6.576709
Skewness	0.021997	-0.125837
Kurtosis	2.260897	2.237055
Jarque-Bera	0.685261	0.806780
Probability	0.709900	0.668051
Sum	451.2600	448.8800
Sum Sq. Dev.	1513.168	1254.340
Observations	30	30
Cross sections	3	3

- o Dans « Data organization » (Cfr figure ci-dessus), Cocher « Cross section specific » pour obtenir les statistiques par individu/firme pour toutes les périodes comme suit :

	I FIRME1	I FIRME2	I FIRME3	P FIRME1	P FIRME2	P FIRME3
Mean	15.50200	15.41500	14.20900	15.39700	16.55900	12.93200
Median	14.93500	17.74000	14.72500	14.60000	17.51000	12.04500
Maximum	29.82000	20.30000	26.68000	26.45000	22.93000	24.97000
Minimum	2.620000	7.630000	3.870000	5.430000	9.160000	1.470000
Std. Dev.	8.448096	4.825251	8.503865	7.167053	4.527863	7.738222
Skewness	0.172028	-0.563074	0.076323	0.162540	-0.347907	0.147495
Kurtosis	2.353563	1.662029	1.533531	2.158319	1.925079	1.836595
Jarque-Bera	0.223440	1.274323	0.905764	0.339210	0.683172	0.600220
Probability	0.894295	0.528791	0.635793	0.843998	0.710642	0.740737
Sum	155.0200	154.1500	142.0900	153.9700	165.5900	129.3200
Sum Sq. Dev.	642.3330	209.5475	650.8415	462.2998	184.5139	538.9208
Observations	10	10	10	10	10	10

- o Dans « Data organization » (Cfr figure ci-haut), Cocher « Time period specific » pour obtenir les statistiques par période pour tous les individus/firmes comme suit :

obs	Mean I?	Med I?	Sd I?	Min I?	Max I?	Mean P?	Med P?	Sd P?	Min P?	Max P?
1982	14.15667	13.32000	5.770670	8.850000	20.30000	14.81000	12.85000	7.338992	8.650000	22.93000
1983	20.57667	17.96000	4.962603	17.47000	26.30000	20.06667	17.96000	4.920715	16.55000	25.69000
1984	5.266667	3.870000	3.556970	2.620000	9.310000	5.370000	5.480000	3.846180	1.470000	9.160000
1985	19.04667	18.01000	4.711330	14.94000	24.19000	19.16333	18.73000	5.602583	13.79000	24.97000
1986	9.140000	7.630000	6.048066	3.990000	15.80000	10.57667	11.31000	5.238638	5.010000	15.41000
1987	12.59000	12.20000	7.063080	5.730000	19.84000	14.01333	12.59000	6.542173	8.300000	21.15000
1988	18.45667	14.93000	7.145602	13.76000	26.68000	18.49000	16.64000	3.654873	16.13000	22.70000
1989	17.10333	11.49000	11.03813	10.00000	29.82000	15.47333	11.61000	9.643964	8.360000	26.45000
1990	19.44000	19.51000	0.917006	18.49000	20.32000	18.21000	19.55000	2.399312	15.44000	19.64000
1991	14.64333	18.32000	8.642895	4.770000	20.84000	13.45333	17.06000	6.960204	5.430000	17.87000
1992	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1993	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

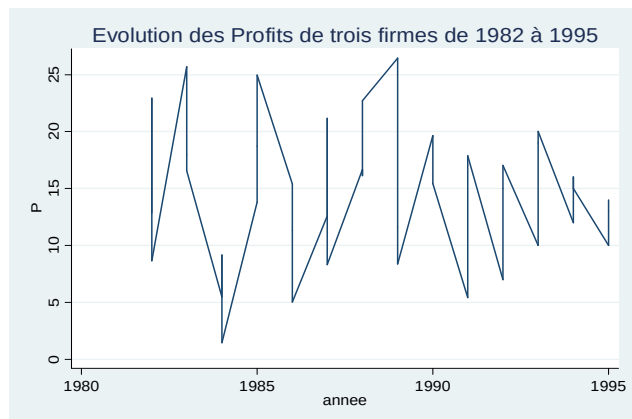
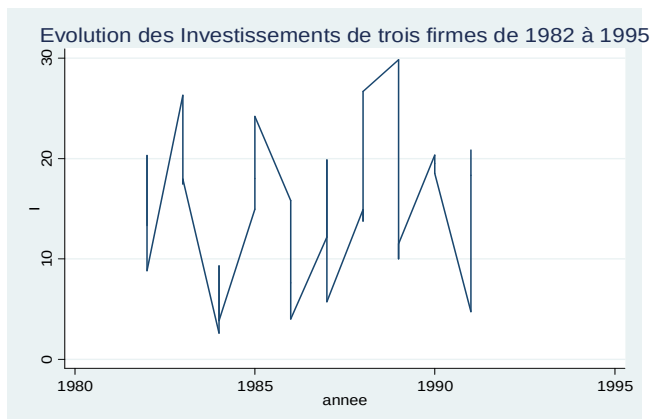


► Evolutions graphiques des variables (I et P)

Sur EViews, pas des commandes pour visualiser graphiquement les variables empilées (seul le graphique par individu/firme pour toutes les périodes est disponible).

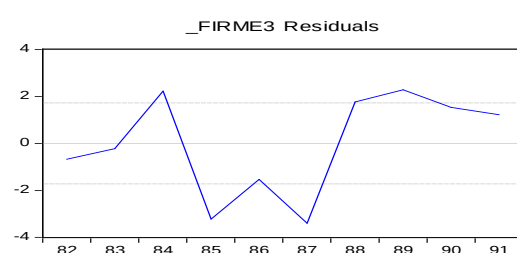
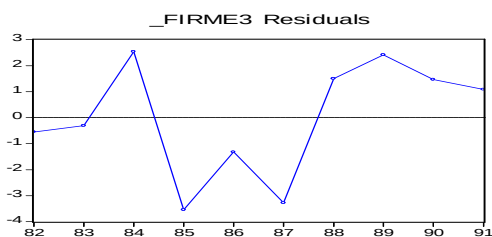
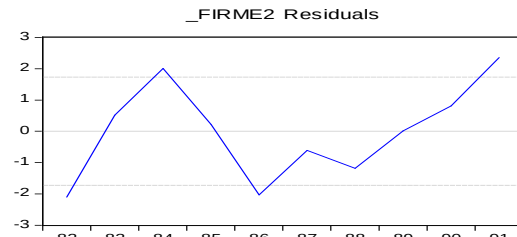
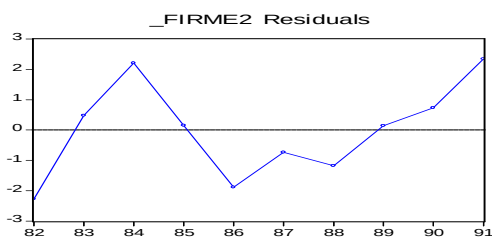
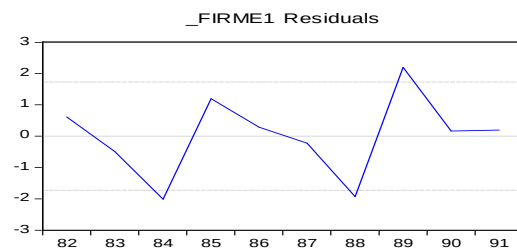
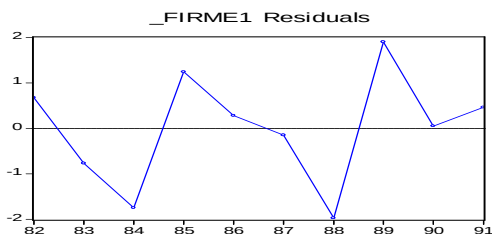
Sur Stata, faire (constat : les séries semblent stationnaires) :

- o Graphique à gauche : `twoway (line I annee), title(Evolution des Investissements de trois firmes de 1982 à 1995)`
- o Graphique à droite : `twoway (line P annee), title(Evolution des Profits de trois firmes de 1982 à 1995)`



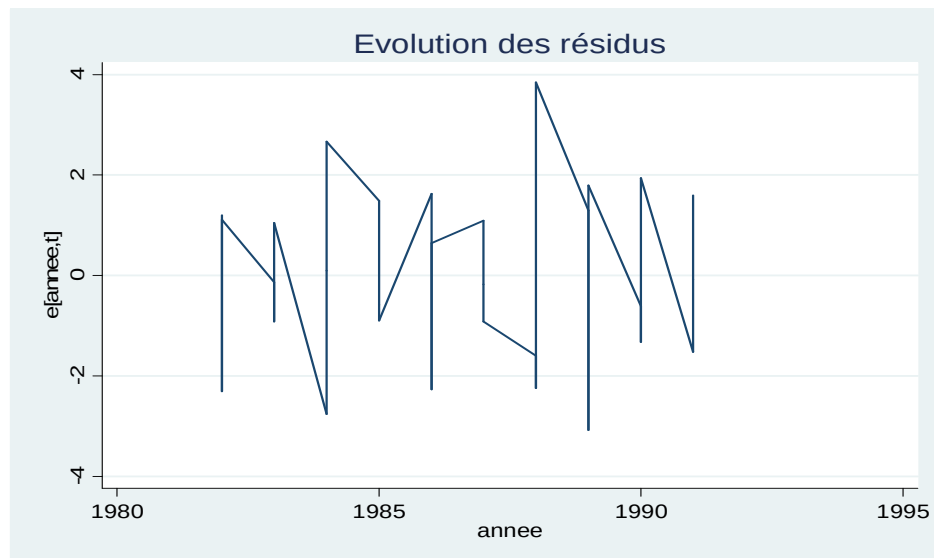
► Evolutions graphiques des résidus

Sur EViews, dans l'output des résultats, cliquer : View/Residuals/Graphs : (à gauche : résultat sur EViews 5.1 et, à droite : résultat sur EViews 7) :



_____ Sur Stata, taper (graphique des résidus pour de données empilées) :

```
{ xtreg I P, fe
  predict res, e
  line res annee, title(Evolution des résidus)
```



CHAP VI LES MODELES A VARIABLES QUALITATIVES

Cas pratique 6.1 : Créer les variables qualitatives sur base d'autres variables

► Travail demandé :

- o Créer/générer la variable : $X = \begin{cases} 1, \text{ si } AGE < 37 \\ 0, \text{ ailleurs} \end{cases}$;
- o Créer/générer la variable : $Y = \begin{cases} 2, \text{ si } AGE < 36 \\ 1, \text{ si } 36 \leq AGE < 50 \\ 0, \text{ si } AGE \geq 50 \end{cases}$;

► Procédures à suivre/Commandes (Sur Eviews) :

_____ Sur EViews :

- o Pour le 1^{er} cas, taper : GENR X=1*(AGE<37)+0*(AGE>=37) ;
- o Pour le 2^{ème} cas, taper :
GENR Y=2*(AGE<36)+1*((AGE>=36) AND (AGE<50))+0*(AGE>=50)

_____ Sur Stata :

- o Pour le 1^{er} cas, taper :

```
{ gen X=0
  replace X=1 if AGE<37
```



o Pour le 2^{ème} cas, taper :

```
{ gen Y=0
  replace Y=2 if AGE<36
  replace Y=1 if AGE>=36&AGE<50
```

_____ Les résultats (variables générées) se présentent comme suit :

X	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0
Y	2	1	1	1	0	1	0	1	2	2	1	0	1	0	1

Cas pratique 6.2 : Relation (lien) entre les variables qualitatives

► Données :

- Taille de l'échantillon (N) : 15 personnes ;
- Variable dépendante $ABON = \begin{cases} 1 = \text{la personne veut s'abonner} \\ 0 = \text{la personne ne veut pas s'abonner} \end{cases}$
- Variables explicatives :
 - o $SEXE = \begin{cases} 1 = \text{Feminin} \\ 0 = \text{Masculin} \end{cases}$;
 - o $AGE = \text{quantitative (discrète)}$

► Travail Demandé :

- Trouver la proportion de personnes qui veulent s'abonner et celle des femmes dans l'échantillon ;
- Utiliser le test de Khi-deux pour vérifier si le « SEXE » joue significativement sur la décision (volonté ou pas) de s'abonner ;
- Considérant la variable « SEXE », générer les modalités « Féminin et Masculin » ;
- Estimer la probabilité de s'abonner (« ABON ») à l'aide du modèle/procédure « LOGIT », les variables « SEXE » et « AGE » étant explicatives, et commenter les résultats obtenus ;
- Vérifier la qualité du modèle LOGIT estimé ou la bonté de l'ajustement par le test de Hosmer-Lemeshow ;
- Calculer le pourcentage de bonnes prédictions et celui de prédictions fausses ;
- Estimer de nouveau la probabilité de s'abonner (« ABON »), cette fois-ci à l'aide des modèles « PROBIT » et « GOMBIT », les variables « SEXE » et « AGE » étant explicatives. Aussi, pour chaque modèle, calculer le pourcentage de prédictions fausses.

► Résolutions :

a) Proportion de personnes qui veulent s'abonner (ABON) et celle des femmes dans l'échantillon (SEXE)

_____ Sur EViews, faire :

```
{ create u 1 15
  data AGE SEXE ABON
```



→ Dans l'output des données relatif aux variables « SEXE » et « ABON » (faire : show SEXE ABON), suivre : *View/N-Way Tabulation...* → la boîte de dialogue suivante apparaît (elle complète la procédure : dans « Output », cocher « Count ») :

Cliquer sur « OK » pour obtenir les résultats suivants :

Tabulation of SEXE and ABON			
Date: 12/29/13 Time: 20:07			
Sample: 1 15			
Included observations: 15			
Tabulation Summary			
<u>Variable</u>	<u>Categories</u>		
SEXE	2		
ABON	2		
Product of Categories	4		
<u>Count</u>	<u>ABON</u>		
	0	1	Total
SEXE	0	1	5
	1	6	10
Total	8	7	15

Sur Stata, faire : `tabulate SEXE ABON, row`

+-----+			
Key			

frequency			
row percentage			
+-----+			
SEXE	ABON		Total
	0	1	
0	4	1	5
	80.00	20.00	100.00
1	4	6	10
	40.00	60.00	100.00
Total	8	7	15
	53.33	46.67	100.00



_____ **Calcul des proportions** : Proportion de personnes voulant s'abonner (PA) et Proportion de femmes dans l'échantillon (PF) :

- o $PA = 7/15 = 0.4667$;
- o $PF = 10/15 = 0.6667$.

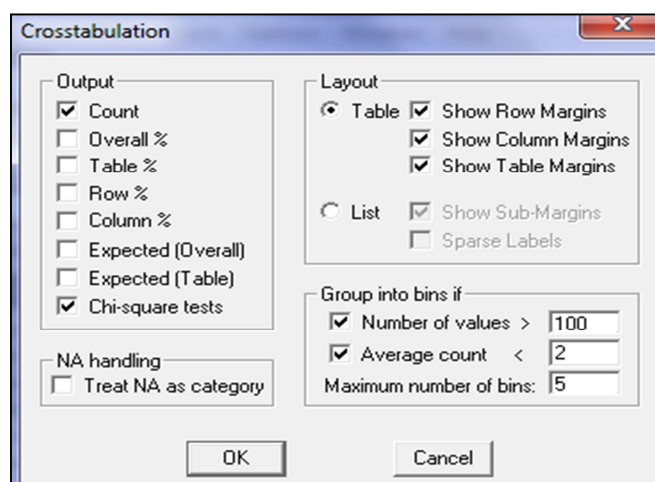
b) Test de Khi-deux pour vérifier si le « SEXE » joue significativement sur la décision (volonté ou pas) de s'abonner

Les hypothèses du test d'indépendance de Khi-deux sont :

H_0 : Egalité des paramètres/Les caractères sont indépendants ($prob > 5\%$)

H_1 : Non égalité des paramètres/Les caractères sont dépendants ($prob < 5\%$)

_____ **Sur EViews**, dans l'output des données relatif aux variables « SEXE » et « ABON » (faire : show SEXE ABON), suivre : View/N-Way Tabulation... → la boîte de dialogue suivante apparaît (elle complète la procédure : dans « Output », cocher « Count » et « Chi-square tests ») :



Cliquer sur « OK » pour obtenir les résultats ci-dessous :

Tabulation of SEXE and ABON			
Date: 12/29/13 Time: 20:22			
Sample: 1 15			
Included observations: 15			
Tabulation Summary			
<u>Variable</u>	<u>Categories</u>		
SEXE	2		
ABON	2		
Product of Categories	4		
<u>Measures of Association</u>	<u>Value</u>		
Phi Coefficient	0.377964		
Cramer's V	0.377964		
Contingency Coefficient	0.353553		
<u>Test Statistics</u>	<u>df</u>	<u>Value</u>	<u>Prob</u>
Pearson X2	1	2.142857	0.1432
Likelihood Ratio G2	1	2.263442	0.1325
WARNING: Expected value is less than 5 in 75.00% of cells (3 of 4).			
<u>Count</u>		<u>ABON</u>	
		0	1
SEXE	0	4	1
	1	4	6
Total		8	7
			Total
			5
			10
			15



_____ **Sur Stata**, faire : `tab ABON SEXE, chi2`

ABON	SEXE		Total
	0	1	
0	4	4	8
1	1	6	7
Total	5	10	15
Pearson chi2(1) = 2.1429 Pr = 0.143			

_____ **Commentaire** : La statistique de khi-deux calculée (Cfr « Pearson X2 » sur le tableau de résultats) est égale à « 2.1429 » et la probabilité associée est de « 0.1432 » (soit, prob>5%). Ainsi, il y a lieu d'accepter l'hypothèse nulle/Ho selon laquelle les variables/caractères « SEXE » et « ABON » ne sont pas liés.

c) Création des modalités « Féminin et Masculin » étant donné la variable « SEXE »

Il est question de générer les modalités :

$$o \text{ } Feminin (FEM) = \begin{cases} 1, si \text{ } SEXE = 1 \\ 0, ailleurs \end{cases};$$

$$o \text{ } Masculin (MSC) = \begin{cases} 1, si \text{ } SEXE = 0 \\ 0, ailleurs \end{cases}$$

_____ **Sur EViews**, faire (respectivement pour les modalités Masculin et Féminin) :

- o `genr MSC=1*(SEXE=0)+0*(SEXE<>0)`
- o `genr FEM=1*(SEXE=1)+0*(SEXE<>1)`

_____ **Sur Stata**, faire :

```
gen MSC=0
replace MSC=1 if SEXE<1
```

```
gen FEM=0
replace FEM=1 if SEXE >0
```

d) Estimation de la probabilité de s'abonner (« ABON ») à l'aide du modèle/procédure « LOGIT », les variables « SEXE » et « AGE » étant explicatives (suivi des commentaires)

_____ NB : la modalité « Féminin » sert de référence.

► **Estimation**

_____ **Sur EViews**, suivre : Quick/Estimate Equation... → dans « Method » (Cfr boîte de dialogue), choisir : BINARY-Binary choice (logit, probit, extreme value) → les boîtes de dialogue suivantes s'affichent/Cfr onglets « Specification » et « Options » (elles complètent la procédure) :



Cliquer sur « **OK** » pour obtenir les résultats qui suivent :

Dependent Variable: ABON
 Method: ML - Binary Logit (Quadratic hill climbing)
 Date: 12/29/13 Time: 21:16
 Sample: 1 15
 Included observations: 15
 Convergence achieved after 6 iterations
 QML (Huber/White) standard errors & covariance

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	14.08807	5.718978	2.463390	0.0138
AGE	-0.288086	0.112841	-2.553038	0.0107
MSC	-2.140715	1.411391	-1.516741	0.1293
Mean dependent var	0.466667	S.D. dependent var		0.516398
S.E. of regression	0.432730	Akaike info criterion		1.250128
Sum squared resid	2.247068	Schwarz criterion		1.391738
Log likelihood	-6.375958	Hannan-Quinn criter.		1.248619
Restr. log likelihood	-10.36385	Avg. log likelihood		-0.425064
LR statistic (2 df)	7.975784	McFadden R-squared		0.384789
Probability(LR stat)	0.018539			
Obs with Dep=0	8	Total obs		15
Obs with Dep=1	7			

Sur Stata, faire : `logit ABON AGE MSC, robust`

```
Iteration 0: log pseudolikelihood = -10.36385
Iteration 1: log pseudolikelihood = -6.8219081
Iteration 2: log pseudolikelihood = -6.4384754
Iteration 3: log pseudolikelihood = -6.3787965
Iteration 4: log pseudolikelihood = -6.3759666
Iteration 5: log pseudolikelihood = -6.3759576

Logistic regression
Log pseudolikelihood = -6.3759576
Number of obs = 15
Wald chi2(2) = 7.43
Prob > chi2 = 0.0244
Pseudo R2 = 0.3848
```

	ABON	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]
	AGE	-.2880861	.1167984	-2.47	0.014	-.5170067 -.0591654
	MSC	-2.140715	1.460923	-1.47	0.143	-5.004072 .7226424
	_cons	14.08807	5.919579	2.38	0.017	2.48591 25.69023

► **Commentaires (explication de certains éléments de l'output) :**



► Les éléments de l'output :

- o « Z-statistic » n'est rien d'autres que le « t de student » calculé ;
- o « ML... » signifie : Maximum Likelihood/Maximum de vraisemblance ;
- o « Mean dependent var et S.D dependent var » : respectivement la moyenne et l'écart-type (Standard Deviation) de la variable dépendante ;
- o « S.E of regression » : la somme des carrés expliqués/Régression ;
- o « Sum squared resid » : Somme des carrés des résidus/Erreurs ;
- o « Log Likelihood » : logarithme de vraisemblance (valeur totale/maximale) ;
- o « Avg. Log likelihood » : moyenne du logarithme de vraisemblance (log likelihood/taille de l'échantillon) ;
- o « Restr. Log likelihood » : logarithme de vraisemblance restreint/contraint (*contrainte/hypothèse : tous les paramètres, constante exclue, sont nuls*) ;
- o « LR statistic » : Log Likelihood Ratio Statistic (Ratio logarithme de vraisemblance) ; et,
- o « McFadden R-squared » : R^2 de McFadden ou Pseudo- R^2 (indice du ratio de maximum de vraisemblance).

► Significativité individuelle des variables : Seule la variable « AGE » influe significativement sur la décision/volonté de s'abonner ($prob < 5\%$) : plus on est âgé/vieux, de moins en moins l'on souhaite s'abonner. Par contre, la modalité « Masculin » – bien que statistiquement non significative – amène à souligner que les femmes ont une forte propension à s'abonner (Sur 7 personnes abonnées, 6 sont des femmes : Cfr tableau de résultats du test de Khi-deux) ;

► Significativité globale des variables/bonté de l'ajustement : Au regard du « LR statistic » – dont la probabilité est $< 5\%$ (soit : 0,018539) – nous concluons à la significativité globale des paramètres. Toutefois, le « R^2 de McFadden » – égal à 38,47% – signale le faible pouvoir explicatif du modèle ;

► « Mean dependent var » = 0.4667 : autant dire qu'il y a 46,67% d'abonnés (ABON=1) dans l'échantillon considéré.

e) Vérification de la qualité du modèle LOGIT estimé ou la bonté de l'ajustement par le test de Hosmer-Lemeshow

Les hypothèses du test d'Hosmer-Lemeshow sont :

H_0 : L'ajustement est bon/Goodness of fit ($prob > 5\%$)

H_1 : L'ajustement n'est pas bon ($prob < 5\%$)

____ Sur EViews (6), dans l'output des résultats, cliquer : View/Goodness-of-Fit Test (Hosmer-Lemeshow) → la boîte de dialogue à gauche apparaît (elle complète la procédure) → cliquer sur « OK » pour obtenir les résultats à droite :



Goodness-of-Fit Test

Form cells based upon

☒ Predicted risk (Hosmer-Lemeshow test)

☐ Series or series expression:

Group observations by

☒ Quantiles 10

☒ Randomize ties to balance cell sizes

☐ Distinct values

Maximum # of cells: 100

Goodness-of-Fit Evaluation for Binary Specification

Andrews and Hosmer-Lemeshow Tests

Equation: UNTITLED

Date: 12/30/13 Time: 18:51

Grouping based upon predicted risk (randomize ties)

	Quantile of Risk		Dep=0		Dep=1		Total Obs	H-L Value
	Low	High	Actual	Expect	Actual	Expect		
1	0.0264	0.0264	1	0.97364	0	0.02636	1	0.02707
2	0.0517	0.1146	2	1.83371	0	0.16629	2	0.18137
3	0.1690	0.1690	1	0.83099	0	0.16901	1	0.20338
4	0.2134	0.2657	2	1.52090	0	0.47910	2	0.63002
5	0.3255	0.3255	0	0.67446	1	0.32554	1	2.07183
6	0.3534	0.5646	0	1.08202	2	0.91798	2	2.35740
7	0.6337	0.6337	1	0.36632	0	0.63368	1	1.72983
8	0.6337	0.6976	1	0.66868	1	1.33132	2	0.24662
9	0.9821	0.9821	0	0.01790	1	0.98210	1	0.01822
10	0.9821	0.9865	0	0.03137	2	1.96863	2	0.03187
Total			8	8.00000	7	7.00000	15	7.49763

H-L Statistic	7.4976	Prob. Chi-Sq(8)	0.4840
Andrews Statistic	12.7923	Prob. Chi-Sq(10)	0.2355

Sur Stata, faire :

Logistic model for ABON, **goodness-of-fit test**

```

number of observations =      15
number of covariate patterns =    13
Pearson chi2(10) =      8.23
Prob > chi2 =    0.6062

```

Commentaire : la probabilité associée à la statistique de Hosmer-Lemeshow/H-L Statistic calculée étant $> 5\%$ (soit, prob = 0.4840), il y a lieu considérer que l'ajustement est bon. **NB** : plus « H-L stat » est grande, de plus en plus la différence entre les valeurs observées et prédites (des déciles : quantile égal à 10) est prononcée, avec comme conséquence le rejet de l'hypothèse nulle.

f) Calcul du pourcentage de bonnes prédictions et celui de prédictions fausses

Sur Stata, faire : `lstat`

```

Logistic model for ABON
----- True -----
Classified |      D      ~D      Total
-----+-----+-----
      +      |      5       2       7
      -      |      2       6       8
-----+-----+-----
    Total   |      7       8      15
Classified + if predicted Pr(D) >= .5
True D defined as ABON != 0
-----+-----+-----
Sensitivity      Pr( + | D)    71.43%
Specificity      Pr( - | ~D)   75.00%
Positive predictive value  Pr( D | +)    71.43%
Negative predictive value  Pr( ~D | -)   75.00%
-----+-----+-----
False + rate for true ~D    Pr( + | ~D)   25.00%
False - rate for true D     Pr( - | D)    28.57%
False + rate for classified + Pr( ~D | +)   28.57%
False - rate for classified - Pr( D | -)    25.00%
-----+-----+-----
Correctly classified              73.33%
-----+-----+-----

```



Sur **EViews**, dans l'output des résultats, cliquer : View/Prediction-Expectation Evaluation→la boîte de dialogue à gauche apparaît (prediction evaluation: 0.5)→cliquer sur « OK » pour obtenir les résultats à droite :

Prediction evaluation

Success if probability is greater than:

Expectation-Prediction Evaluation for Binary Specification
Equation: UNTITLED
Date: 12/30/13 Time: 20:14
Success cutoff: C = 0.5

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)≤C	6	2	8	8	7	15
P(Dep=1)>C	2	5	7	0	0	0
Total	8	7	15	8	7	15
Correct	6	5	11	8	0	8
% Correct	75.00	71.43	73.33	100.00	0.00	53.33
% Incorrect	25.00	28.57	26.67	0.00	100.00	46.67
Total Gain*	-25.00	71.43	20.00			
Percent Gain**	NA	71.43	42.86			

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
E(# of Dep=0)	5.83	2.17	8.00	4.27	3.73	8.00
E(# of Dep=1)	2.17	4.83	7.00	3.73	3.27	7.00
Total	8.00	7.00	15.00	8.00	7.00	15.00
Correct	5.83	4.83	10.66	4.27	3.27	7.53
% Correct	72.85	68.97	71.04	53.33	46.67	50.22
% Incorrect	27.15	31.03	28.96	46.67	53.33	49.78
Total Gain*	19.52	22.30	20.82			
Percent Gain**	41.82	41.82	41.82			

*Change in "% Correct" from default (constant probability) specification
**Percent of incorrect (default) prediction corrected by equation

Commentaires :

- o Le taux de bonnes prédictions (TBP) = 73.33% [$TBP = (6 + 5)/15 * 100$] : prendre les éléments sur la diagonale principale (TBP > 50% : l'ajustement est bon) ;
 - o Le taux de prédictions fausses (TFP) = 26.67% [$TFP = (2 + 2)/15 * 100$] : prendre les éléments équidistants à la diagonale principale (TFP < 50% : l'ajustement est bon).
- g) **Estimation de la probabilité de s'abonner (« ABON ») à l'aide des modèles « PROBIT » et « GOMBIT », les variables « SEXE » et « AGE » étant explicatives (pour chaque modèle, calcul du pourcentage de prédictions fausses)**

NB : la modalité « Feminin » sert de référence.

► Estimation du modèle PROBIT

Sur **EViews** (6), suivre : Quick/Estimate Equation... → dans « Method » (Cfr boîte de dialogue), choisir : BINARY-Binary choice (logit, probit, extreme value) → [dans l'onglet « Specification » : cocher « Probit » ; et dans l'onglet « Options » : cocher « Huber/White »] → OK :



Dependent Variable: ABON					
Method: ML - Binary Probit (Quadratic hill climbing)					
Date: 12/30/13 Time: 21:08					
Sample: 1 15					
Included observations: 15					
Convergence achieved after 5 iterations					
QML (Huber/White) standard errors & covariance					
Variable	Coefficient	Std. Error	z-Statistic	Prob.	
C	8.439950	3.017240	2.797241	0.0052	
AGE	-0.172894	0.059294	-2.915894	0.0035	
MSC	-1.306470	0.818055	-1.597043	0.1103	
McFadden R-squared	0.394091	Mean dependent var			0.466667
S.D. dependent var	0.516398	S.E. of regression			0.431277
Akaike info criterion	1.237273	Sum squared resid			2.232001
Schwarz criterion	1.378883	Log likelihood			-6.279545
Hannan-Quinn criter.	1.235764	Restr. log likelihood			-10.36385
LR statistic	8.168609	Avg. log likelihood			-0.418636
Prob(LR statistic)	0.016835				
Obs with Dep=0	8	Total obs	15		
Obs with Dep=1	7				

Sur Stata, faire : probit ABON AGE MSC, robust

```

Iteration 0: log pseudolikelihood = -10.36385
Iteration 1: log pseudolikelihood = -6.7419163
Iteration 2: log pseudolikelihood = -6.331693
Iteration 3: log pseudolikelihood = -6.2815948
Iteration 4: log pseudolikelihood = -6.2795503
Iteration 5: log pseudolikelihood = -6.279545

Probit regression
Log pseudolikelihood = -6.279545
Number of obs = 15
Wald chi2(2) = 9.49
Prob > chi2 = 0.0087
Pseudo R2 = 0.3941

```

ABON	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]
AGE	-.1728942	.0613738	-2.82	0.005	-.2931846 -.0526038
MSC	-1.30647	.8467655	-1.54	0.123	-2.9661 .3531601
_cons	8.43995	3.123089	2.70	0.007	2.318808 14.56109

Pourcentage de prédictions fausses (TFP): Cfr « modèle Logit » pour la procédure/commandes (Sur Stata ou EViews).

Expectation-Prediction Evaluation for Binary Specification						
Equation: UNTITLED						
Date: 12/30/13 Time: 21:24						
Success cutoff: C = 0.5						
	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)<=C	6	2	8	8	7	15
P(Dep=1)>C	2	5	7	0	0	0
Total	8	7	15	8	7	15
Correct	6	5	11	8	0	8
% Correct	75.00	71.43	73.33	100.00	0.00	53.33
% Incorrect	25.00	28.57	26.67	0.00	100.00	46.67
Total Gain*	-25.00	71.43	20.00			
Percent Gain**	NA	71.43	42.86			
	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
E(# of Dep=0)	5.90	2.17	8.08	4.27	3.73	8.00
E(# of Dep=1)	2.10	4.83	6.92	3.73	3.27	7.00
Total	8.00	7.00	15.00	8.00	7.00	15.00
Correct	5.90	4.83	10.73	4.27	3.27	7.53
% Correct	73.80	68.97	71.55	53.33	46.67	50.22
% Incorrect	26.20	31.03	28.45	46.67	53.33	49.78
Total Gain*	20.47	22.30	21.33			
Percent Gain**	43.87	41.82	42.84			
*Change in "% Correct" from default (constant probability) specification						
**Percent of incorrect (default) prediction corrected by equation						



Le taux de prédictions fausses (TFP) = 26.67% [$TFP = (2 + 2)/15 * 100$] : prendre les éléments équidistants à la diagonale principale ($TFP < 50\%$: l'ajustement est bon).

► **Estimation du modèle GOMBIT**

_____ **Sur EViews (6)**, suivre : Quick/Estimate Equation... → dans « Method » (Cfr boîte de dialogue), choisir : BINARY-Binary choice (logit, probit, extreme value) → dans l'onglet « Specification » : cocher « **Extreme value** » ; et dans l'onglet « Options » : cocher « Huber/White » → OK :

Dependent Variable: ABON				
Method: ML - Binary Extreme Value (Quadratic hill climbing)				
Date: 12/30/13 Time: 21:26				
Sample: 1 15				
Included observations: 15				
Convergence achieved after 5 iterations				
QML (Huber/White) standard errors & covariance				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	11.16550	4.780393	2.335687	0.0195
AGE	-0.217495	0.088311	-2.462844	0.0138
MSC	-1.707759	1.040990	-1.640515	0.1009
McFadden R-squared	0.409186	Mean dependent var	0.466667	
S.D. dependent var	0.516398	S.E. of regression	0.425981	
Akaike info criterion	1.216414	Sum squared resid	2.177517	
Schwarz criterion	1.358024	Log likelihood	-6.123105	
Hannan-Quinn criter.	1.214906	Restr. log likelihood	-10.36385	
LR statistic	8.481490	Avg. log likelihood	-0.408207	
Prob(LR statistic)	0.014397			
Obs with Dep=0	8	Total obs	15	
Obs with Dep=1	7			

_____ **Pourcentage de prédictions fausses (TFP)** : Cfr « modèle Logit » pour la procédure/commandes (sur EViews).

Expectation-Prediction Evaluation for Binary Specification						
Equation: UNTITLED						
Date: 12/30/13 Time: 21:29						
Success cutoff: C = 0.5						
	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)≤C	6	2	8	8	7	15
P(Dep=1)>C	2	5	7	0	0	0
Total	8	7	15	8	7	15
Correct	6	5	11	8	0	8
% Correct	75.00	71.43	73.33	100.00	0.00	53.33
% Incorrect	25.00	28.57	26.67	0.00	100.00	46.67
Total Gain*	-25.00	71.43	20.00			
Percent Gain**	NA	71.43	42.86			
	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
E(# of Dep=0)	5.98	2.06	8.04	4.27	3.73	8.00
E(# of Dep=1)	2.02	4.94	6.96	3.73	3.27	7.00
Total	8.00	7.00	15.00	8.00	7.00	15.00
Correct	5.98	4.94	10.92	4.27	3.27	7.53
% Correct	74.74	70.52	72.77	53.33	46.67	50.22
% Incorrect	25.26	29.48	27.23	46.67	53.33	49.78
Total Gain*	21.41	23.85	22.55			
Percent Gain**	45.88	44.72	45.30			
*Change in "% Correct" from default (constant probability) specification						
**Percent of incorrect (default) prediction corrected by equation						

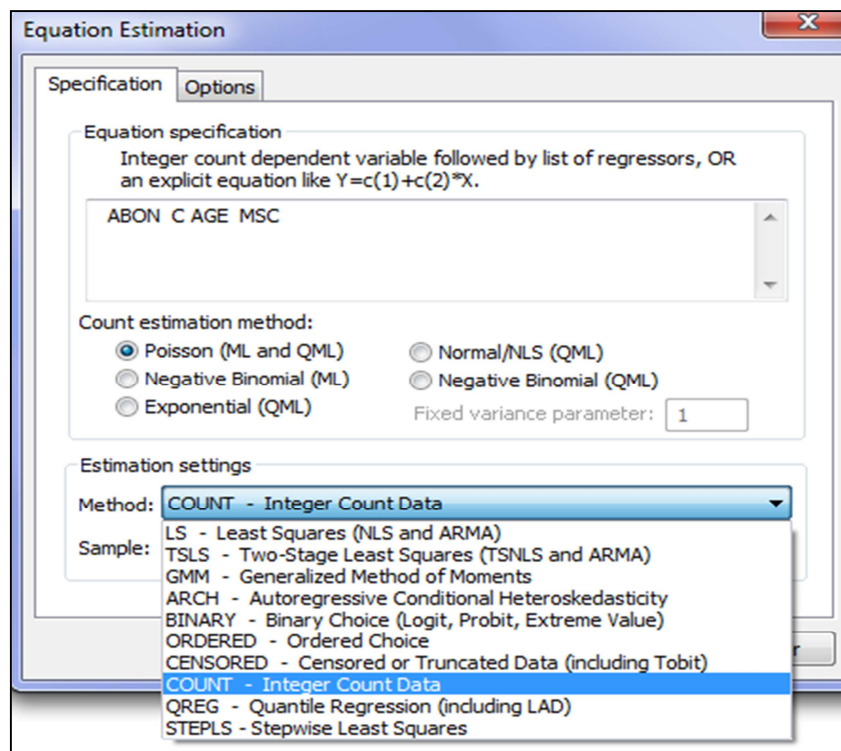


Le taux de prédictions fausses (TFP) = 26.67% [$TFP = (2 + 2)/15 * 100$] : prendre les éléments équidistants à la diagonale principale ($TFP < 50\%$: l'ajustement est bon).

► **Constat** : les résultats d'estimation de nos trois modèles diffèrent très peu.

h) **ANNEXES : Estimation du modèle de Poisson et du modèle Binomial Négatif(NegBin)**

Sur EViews, suivre (la boîte de dialogue ci-dessous illustre la procédure) :



o Cocher « Poisson (ML and QML) » → OK :

Dependent Variable: ABON				
Method: ML/QML - Poisson Count (Quadratic hill climbing)				
Date: 12/30/13 Time: 22:05				
Sample: 1 15				
Included observations: 15				
Convergence achieved after 5 iterations				
Covariance matrix computed using second derivatives				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	2.461433	2.181419	1.128363	0.2592
AGE	-0.068359	0.051941	-1.316092	0.1881
MSC	-0.870811	1.113499	-0.782049	0.4342
R-squared	0.371984	Mean dependent var		0.466667
Adjusted R-squared	0.267315	S.D. dependent var		0.516398
S.E. of regression	0.442021	Akaike info criterion		1.831794
Sum squared resid	2.344591	Schwarz criterion		1.973404
Log likelihood	-10.73845	Hannan-Quinn criter.		1.830285
Restr. log likelihood	-12.33498	LR statistic		3.193054
Avg. log likelihood	-0.715897	Prob(LR statistic)		0.202599



o Cocher « Negative Binomial (QML) » et « Negative Binomial (ML) » → OK :

Dependent Variable: ABON Method: QML - Negative Binomial Count (Quadratic hill climbing) Date: 12/30/13 Time: 22:08 Sample: 1 15 Included observations: 15 QML parameter used in estimation: 1 Convergence achieved after 5 iterations Covariance matrix computed using second derivatives					Dependent Variable: ABON Method: ML - Negative Binomial Count (Quadratic hill climbing) Date: 12/30/13 Time: 22:29 Sample: 1 15 Included observations: 15 Convergence achieved after 11 iterations WARNING: Singular covariance - coefficients are not unique Covariance matrix computed using second derivatives				
Variable	Coefficient	Std. Error	z-Statistic	Prob.	Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	2.755213	2.957921	0.931469	0.3516	C	2.461433	NA	NA	NA
AGE	-0.074839	0.067557	-1.107788	0.2680	AGE	-0.068359	NA	NA	NA
MSC	-0.893460	1.241191	-0.719840	0.4716	MSC	-0.870811	NA	NA	NA
R-squared 0.353273 Mean dependent var 0.466667 Adjusted R-squared 0.245486 S.D. dependent var 0.516398 S.E. of regression 0.448558 Akaike info criterion 2.083552 Sum squared resid 2.414446 Schwarz criterion 2.225162 Log likelihood -12.62664 Hannan-Quinn criter. 2.082044 Restr. log likelihood -13.76081 LR statistic 2.268340 Avg. log likelihood -0.841776 Prob(LR statistic) 0.321689					Mixture Parameter SHAPE:C(4) -16.36914 NA NA NA				
R-squared 0.371984 Mean dependent var 0.466667 Adjusted R-squared 0.200707 S.D. dependent var 0.516398 S.E. of regression 0.461676 Akaike info criterion 1.965127 Sum squared resid 2.344591 Schwarz criterion 2.153941 Log likelihood -10.73845 Hannan-Quinn criter. 1.963116 Restr. log likelihood -12.33498 LR statistic 3.193054 Avg. log likelihood -0.715897 Prob(LR statistic) 0.362807									

Sur Stata, taper : nbreg ABON AGE MSC

Fitting Poisson model:						
Iteration 0:	log likelihood =	-10.738555				
Iteration 1:	log likelihood =	-10.738453				
Iteration 2:	log likelihood =	-10.738453				
Fitting constant-only model:						
Iteration 0:	log likelihood =	-13.76081				
Iteration 1:	log likelihood =	-12.33498				
Iteration 2:	log likelihood =	-12.33498	(not concave)			
Fitting full model:						
Iteration 0:	log likelihood =	-10.937678				
Iteration 1:	log likelihood =	-10.738458				
Iteration 2:	log likelihood =	-10.738453				
Iteration 3:	log likelihood =	-10.738453	(not concave)			
Negative binomial regression						
Dispersion = mean				Number of obs =	15	
Log likelihood = -10.738453				LR chi2(2) =	3.19	
				Prob > chi2 =	0.2026	
				Pseudo R2 =	0.1294	
ABON	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
AGE	-.0683594	.0519412	-1.32	0.188	-.1701624	.0334435
MSC	-.8708107	1.113499	-0.78	0.434	-3.053229	1.311608
_cons	2.461433	2.181419	1.13	0.259	-1.81407	6.736936
/lnalpha	-21.7867	.			.	.
alpha	3.45e-10	.			.	.
Likelihood-ratio test of alpha=0: chibar2(01) = 0.00 Prob>=chibar2 = 1.000						

A savoir :

- o Log likelihood (modèle de poisson) = -10,73845



- o Log likelihood (modèle Binomial négatif) = $-12,62664$
- o L'on compare le modèle de Poisson au modèle Binomial négatif pour tester sa bonté (son caractère réaliste). Pour notre cas illustratif, Le test est celui de la « raison de vraisemblance/LR-Stat » calculée comme suit :

$LR\text{-stat} = 2[\text{Log likelihood (modèle Binomial négatif)} - \text{Log likelihood (modèle de poisson)}] \sim \chi^2_{dl=1}$

Les hypothèses du test :

H_0 : La spécification de Poisson est réaliste/ $LR\text{-stat} < \chi^2_{dl=1}(\text{table})/(\text{prob} > 5\%)$

H_1 : La spécification de Poisson n'est pas réaliste/ $LR\text{-stat} > \chi^2_{dl=1}(\text{table})/(\text{prob} < 5\%)$

$$LR - stat = 2(12,62664 - 10,73845) = 3,77638$$

- o A 99%, seuil (Chi-carré) = 6,635 : $LR\text{-stat} < \text{Seuil}$: La spécification de Poisson est réaliste (la retenir).

CHAP VII

MODELISATION HETEROSCEDASTIQUE : LES MODELES « ARCH » et GARCH

Cas pratique 7.1 : Simulation des processus ARMA

- **Travail demandé :** Simuler (générer) les processus ci-après (avec et : bruit blanc normal centré réduit) :

- o $MA(1) : X_t = 2 + e_t + 0,8 e_{t-1}$
- o $AR(1) : X_t = 2 + 0,9 X_{t-1} + e_t$
- o $MA(2) : X_t = 2 + e_t + 0,6 e_{t-1} - 0,3 e_{t-2}$
- o $AR(2) : X_t = 2 + 0,9 X_{t-1} - 0,7 X_{t-2} + e_t$
- o $ARMA(1,1) : X_t = 2 + 0,9 X_{t-1} + e_t + 0,8 e_{t-1}$
- o $ARIMA(0,1,0) : X_t = 2 + X_{t-1} + e_t$

- **Simulation des processus :**

a) **Bruit blanc : e_t (désigné par « E »)**

_____Création de la variable « E » (processus bruit blanc) :

Sur EViews :

create u 1 200
gen E=NRND

Sur Stata :

set obs 200
gen u=uniform
gen E=invnorm(u)

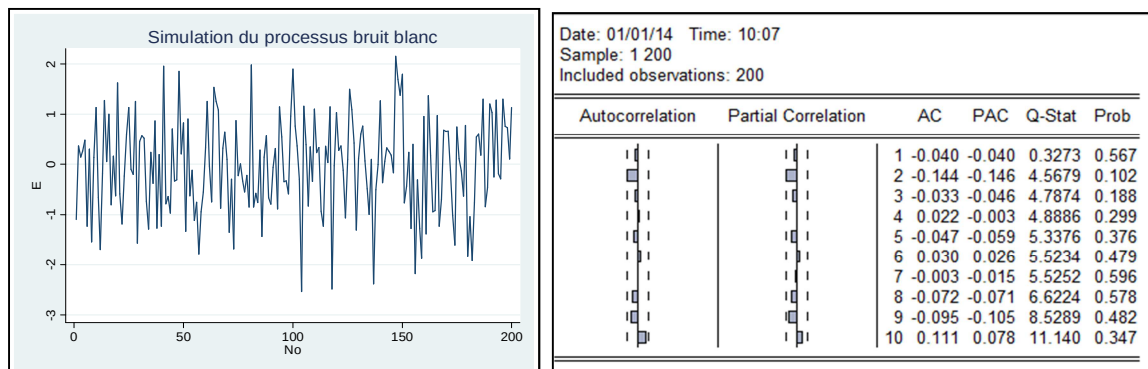
Notes :

- o **E (EViews) :** variable normale centrée réduite ;
- o **E (Stata) :** distribution standard normale centrée sur zéro.



Graphique et corrélogramme du processus bruit blanc :

- * Sur EViews (corrélogramme) : View/Correlogram...→OK ;
- * Sur Stata (Graphique) : après avoir crée la variable “No” (obs : 1 à 200), taper :
line E No, title(Simulation du processus bruit blanc)



- * Caractéristiques : Stationnarité ; pas des pics significatifs (tous les termes sont compris dans l'intervalle de confiance stylisé par des traits horizontaux). NB : AC (fonction d'autocorrélation simple) et PAC (fonction d'autocorrélation partielle).

b) Processus MA(1) : MA1Création de la variable Moyenne Mobile d'ordre 1/MA(1) :**Sur EViews :**

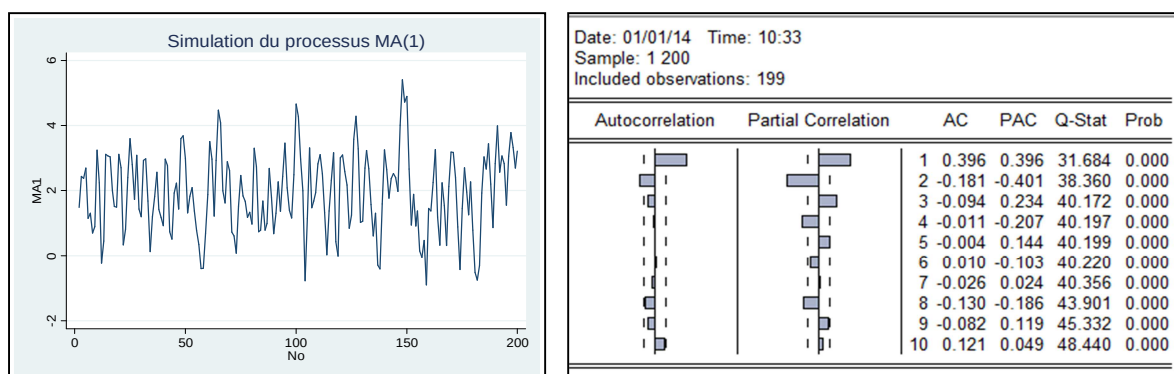
```
create u 1 200
genr MA1=2+E+0.8*E(-1)
```

Sur Stata :

```
set obs 200
gen LE=E[_n-1]
gen MA1=2+E+0.8*LE
```

Graphique et corrélogramme pour MA(1) :

- * Sur EViews (corrélogramme) : View/Correlogram...→OK ;
- * Sur Stata (Graphique) :
line MA1 No, title(Simulation du processus MA(1))



- * Caractéristiques : Stationnarité ; décroissance exponentielle alternée pour la PAC ; et le 1^{er} terme de l'AC est significativement différent de zéro (il sort de l'intervalle de confiance stylisé par des traits horizontaux).



c) Processus AR(1) : AR1

Création du processus Autorégressif d'ordre 1/AR(1) :

Sur EViews :

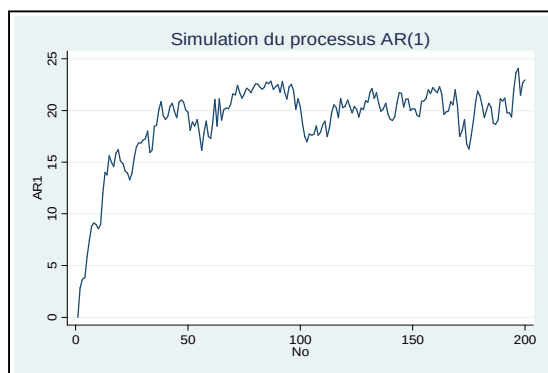
```
smpl 1 1
genr AR1=0
smpl 2 200
genr AR1=2+0.9*AR1(-1)+E
```

Sur Stata :

Copier et coller les données depuis EViews
[variable « AR(1) décalé/AR1[_n-1] »
introuvable au moment de la création de la
variable « AR(1) dans Stata].

Graphique et corrélogramme de AR(1) :

- * Sur EViews (corrélogramme) : View/Correlogram...→OK ;
- * Sur Stata (Graphique) :
line AR1 No, title(Simulation du processus AR(1))



Date: 01/01/14 Time: 16:44 Sample: 2 200 Included observations: 199					
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.900	0.900	163.62	0.000	
2	0.807	-0.014	295.96	0.000	
3	0.726	0.011	403.63	0.000	
4	0.668	0.074	495.06	0.000	
5	0.612	-0.007	572.38	0.000	
6	0.568	0.037	639.28	0.000	
7	0.524	-0.008	696.54	0.000	
8	0.477	-0.033	744.25	0.000	
9	0.431	-0.016	783.30	0.000	
10	0.391	0.011	815.72	0.000	

- * **Caractéristiques** : Non Stationnarité ; décroissance géométrique pour l'AC ; et le 1^{er} terme de la PAC est significativement différent de zéro (il sort de l'intervalle de confiance stylisé par des traits horizontaux).

d) Processus MA(2) : MA2

Création du processus Moyenne Mobile d'ordre 2 MA(2) :

Sur EViews :

```
smpl 1 200
genr MA2=2+E+0.6*E(-1)-0.3*E(-2)
```

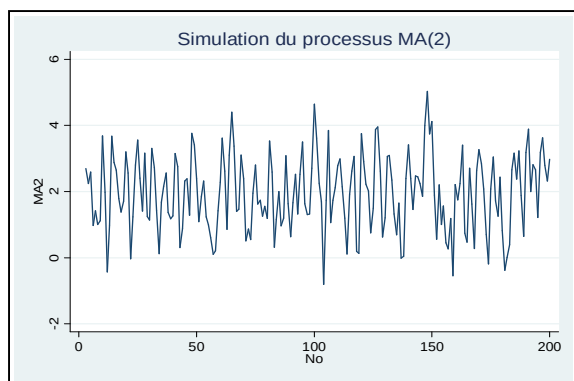
Sur Stata :

```
set obs 200
gen LE=E[_n-1]
gen LE2=E[_n-2]
gen MA2=2+E+0.6*LE-0.3*LE2
```

Graphique et corrélogramme de MA(2) :

- * Sur EViews (corrélogramme) : View/Correlogram...→OK ;
- * Sur Stata (Graphique) :
line MA2 No, title(Simulation du processus MA(2))





Date: 01/01/14 Time: 17:08
Sample: 1 200
Included observations: 198

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.224	0.224	10.060	0.002
		2	-0.352	-0.423	35.082	0.000
		3	-0.048	0.203	35.550	0.000
		4	0.016	-0.244	35.600	0.000
		5	-0.013	0.158	35.637	0.000
		6	0.026	-0.123	35.779	0.000
		7	0.016	0.107	35.834	0.000
		8	-0.126	-0.262	39.152	0.000
		9	-0.103	0.124	41.390	0.000
		10	0.109	-0.062	43.911	0.000

- * **Caractéristiques** : Stationnarité ; décroissance exponentielle alternée pour la PAC ; et les deux premiers termes de l'AC sont significativement différents de zéro (ils sortent de l'intervalle de confiance stylisé par des traits horizontaux : le 1^{er} pic est positif et le second est négatif).

e) Processus AR(2) : AR2

Création du processus Autorégressif d'ordre 2 AR(2) :

Sur EViews :

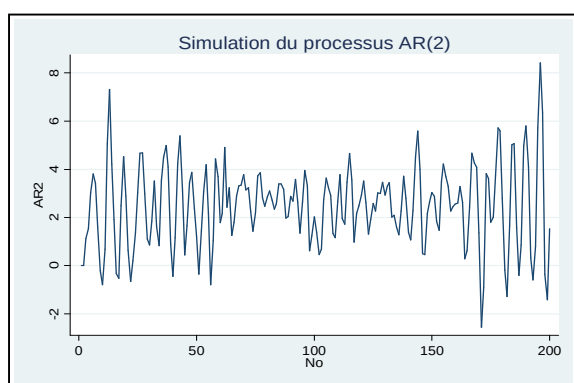
```
smpl 1 2
genr AR2=0
smpl 3 200
genr AR2=2+0.9*AR2(-1)-0.7*AR2(-2)+E
```

Sur Stata :

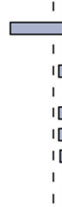
Copier et coller les données depuis EViews [« AR(1) décalé/AR1[_n-1] et [« AR(2) décalé/AR2[_n-2] » introuvables au moment de la création de la variable « AR(2) dans Stata].

Graphique et corrélogramme de AR(2) :

- * Sur EViews (corrélogramme) : View/Correlogram...→OK ;
- * Sur Stata (Graphique) :
line AR2 No, title(Simulation du processus AR(2))



Date: 01/01/14 Time: 17:49
Sample: 3 200
Included observations: 198

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.472	0.472	44.753	0.000
		2 -0.325	-0.704	66.044	0.000
		3 -0.575	0.046	133.21	0.000
		4 -0.245	-0.067	145.48	0.000
		5 0.189	0.012	152.79	0.000
		6 0.294	-0.066	170.58	0.000
		7 0.060	-0.075	171.33	0.000
		8 -0.192	-0.056	179.04	0.000
		9 -0.189	0.011	186.50	0.000
		10 0.046	0.076	186.94	0.000

- * **Caractéristiques** : Stationnarité en moyenne ; décroissance sinusoïdale amortie pour l'AC ; et les deux premiers termes de la PAC sont significativement différents de zéro (ils sortent de l'intervalle de confiance stylisé par des traits horizontaux : le 1^{er} pic est positif et le second est négatif).



f) Processus ARMA(1,1) : ARMA11

_____Création du processus Autorégressif d'ordre 1 et Moyenne Mobile d'ordre 1/
ARMA(1,1) :

Sur EViews :

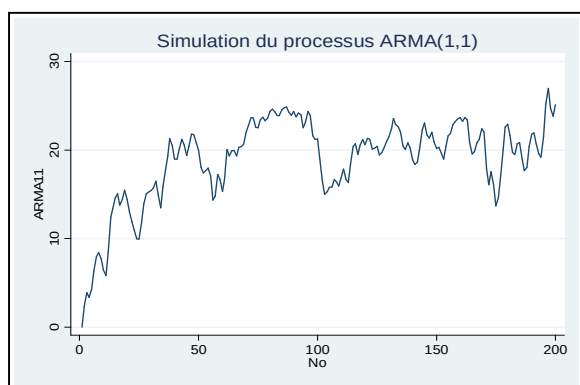
```
smpl 1 1
genr ARMA11=0
smpl 2 200
genr ARMA11=2+0.9*ARMA11(-1)+E+0.8*E(-1)
```

Sur Stata :

Copier et coller les données depuis EViews[« ARMA(1,1)décalé/ARMA11[_n-1] » introuvable au moment de la création de la variable « ARMA11 » dans Stata].

_____Graphique et corrélogramme de ARMA(1,1) :

- * Sur EViews (corrélogramme) : View/Correlogram...→OK ;
- * Sur Stata (Graphique) :
line ARMA11 No, title(Simulation du processus ARMA(1,1))



Date: 01/01/14 Time: 18:08 Sample: 2 200 Included observations: 199					
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.918	0.918	170.16	0.000	
2	0.812	-0.190	304.13	0.000	
3	0.727	0.093	411.93	0.000	
4	0.665	0.068	502.72	0.000	
5	0.609	-0.029	579.16	0.000	
6	0.563	0.056	644.77	0.000	
7	0.521	-0.006	701.25	0.000	
8	0.478	-0.023	749.03	0.000	
9	0.439	0.026	789.64	0.000	
10	0.402	-0.023	823.87	0.000	

- * **Caractéristiques** : Non Stationnarité ; décroissance lente pour l'AC ; et le premier terme de la PAC est significativement différent de zéro (il sort de l'intervalle de confiance stylisé par des traits horizontaux).

g) Processus ARMA Intégré d'ordre 1 : ARIMA/Marche aléatoire

_____Création du processus ARMA intégré d'ordre 1/ARIMA(0,1,0) :

Sur EViews :

```
smpl 1 1
genr ARIMA=0
smpl 2 200
genr ARIMA=2+ARIMA(-1)+E
```

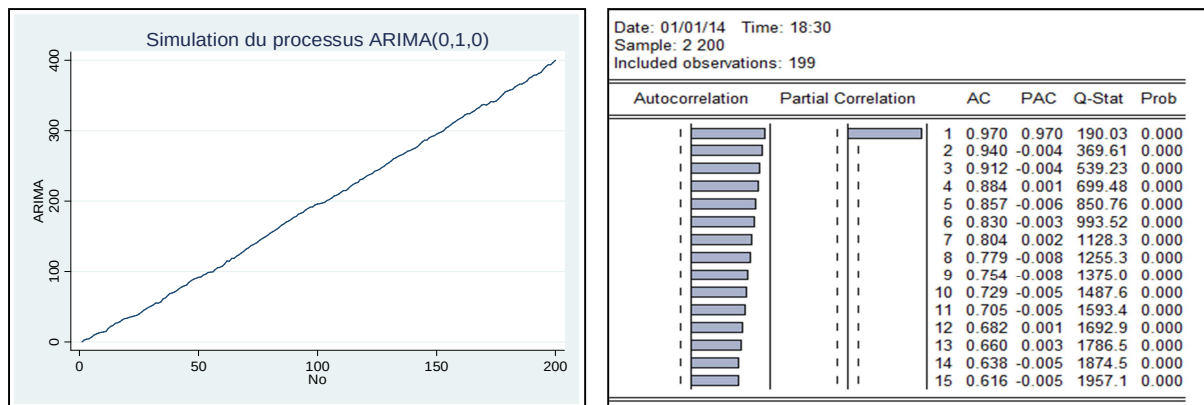
Sur Stata :

Copier et coller les données depuis EViews[« ARIMA(0,1,0)décalé/ARIMA[_n-1] » introuvable au moment de la création de la variable « ARIMA » dans Stata].

_____Graphique et corrélogramme d'ARIMA(1,0,1) ou Marche aléatoire :

- * Sur EViews (corrélogramme) : View/Correlogram...→OK ;
- * Sur Stata (Graphique) :
line ARIMA No, title(Simulation du processus ARIMA(0,1,0))





- * **Caractéristiques:** Non Stationnarité (Processus marche aléatoire affecté d'une tendance); décroissance lente pour l'AC; et le premier terme de la PAC est significativement différent de zéro (il sort de l'intervalle de confiance stylisé par des traits horizontaux).

Cas pratique 7.2 : Simulation d'un processus ARCH(1)

- **Modèle :** Simuler (générer) le processus ARCH(1) ci-après (avec e_t : bruit blanc normal centré réduit) :

$$e_t = u_t \sqrt{1 + 0.8 e_{t-1}^2} \dots \dots [7.0]$$

► **Travail demandé :**

- Générer ce processus (créer la variable EA) ;
- Produire le graphique et le corrélogramme de ce processus ;
- Vérifier la normalité de ce processus (test de Jarque-Berra) ;
- En étudier aussi la linéarité (test de Keenan).

a) Génération du processus ARCH(1) : Modèle (7.1)

Sur EViews :

```
create u 1 200
genr U=NRND
smpl 1 1
genr P=0
smpl 2 200
genr P=1+0.8*U(-1)*U(-1)
genr EA=sqr(P)*U
```

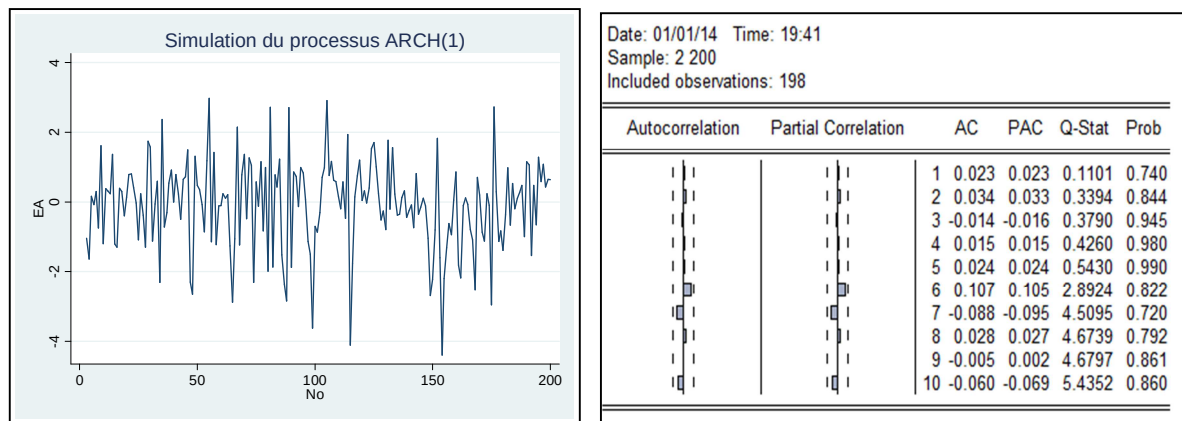
Sur Stata :

```
copier et coller « U » généré sur EViews
gen LU=U[_n-1]
gen P=.
replace P=0 in 1
replace P=1+0.8*LU*LU if No>1
gen EA=sqrt(P)*U
```

b) Graphique et corrélogramme du processus ARCH(1)

- * Sur EViews (corrélogramme) : View/Correlogram...→OK ;
- * Sur Stata (Graphique) :
line EA No, title(Simulation du processus ARCH(1))

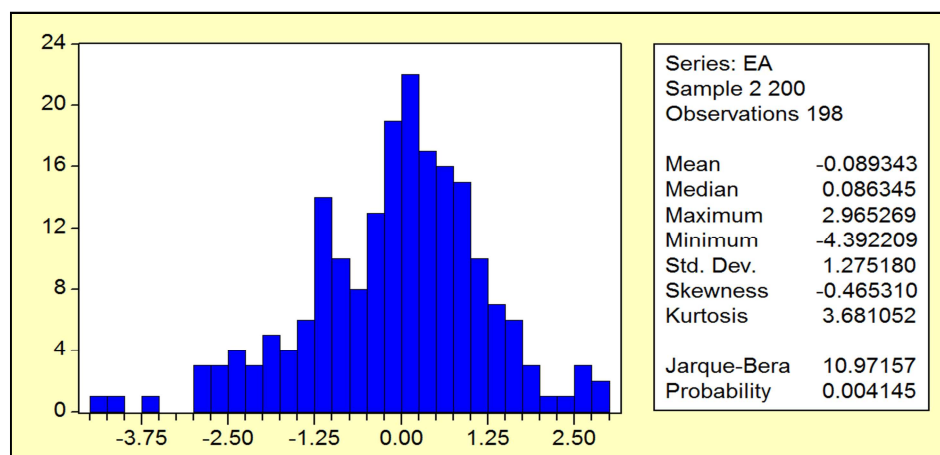




Constat : notre processus ARCH(1) est volatile (forte variabilité pour les observations 80 et 140 à 160), les périodes/moments d'instabilité (variance élevée) faisant régulièrement suite à celles de stabilité.

c) Vérifier la normalité de ce processus (test de Jarque-Berra)

- * Sur Eviews : View/Descriptive Statistics/Histogram and Stats ;
- * Sur Stata : `sktest EA`



Skewness/Kurtosis tests for Normality				
Variable	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	Prob>chi2
EA	0.008	0.068	9.26	0.0098

Commentaires :

- o Skewness (mesure d'asymétrie, il est nul pour une distribution normale) = -0.465 : c.à.d. distribution dissymétrique à gauche (étalée à droite). D'où, présomption de non linéarité (Skewness non nul) ;
- o Kurtosis (mesure d'aplatissement, il avoisine « 3 » pour une distribution normale) = 3.681 > 3 : c.à.d. distribution leptokurtique (centre plus élevé/pointu) ;
- o Jarque-Bera (test de normalité, avec H_0 : le processus est normalement distribué). La probabilité associée est de $0.0041 < 5\%$: rejet de l'hypothèse de



normalité pour le processus ARCH(1) sous-étude (même observation sous Stata).

d) Test de linéarité du processus ARCH(1) de Keenan

Rappels : Pour tester la linéarité d'une série (Y_t), Keenan propose trois étapes :

(i) Estimer le modèle AR(P) suivant :

$$Y_t = \Phi_0 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \dots + \Phi_p Y_{t-p} + u_t$$

(ii) Estimer également :

$$Y_t^2 = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + \dots + a_p Y_{t-p} + e_t$$

(iii) Estimer enfin :

$$\hat{u}_t = \lambda \hat{e}_t + v_t$$

Avec : $t = p + 1, \dots, T$ et p = décalage optimal déterminé à partir des critères d'AKAIKE et SCHWARZ. Ainsi, la statistique « F » du test de Keenan est construit comme suit :

$$\hat{F} = \hat{\beta}^2 \left[\frac{T - 2p - 2}{\sum_{t=p+1}^T \hat{u}_t^2 - \hat{\beta}^2} \right] \sim F_{\alpha(1, T-2p-2)}$$

Avec :

$$\hat{\beta}^2 = \hat{\lambda} \left[\sum_{t=p+1}^T \hat{e}_t^2 \right]^{1/2} = \hat{\lambda}^2 \left[\sum_{t=p+1}^T \hat{e}_t^2 \right]$$

Les hypothèses du test sont :

$H_0 : Y_t$ est linéaire ($\hat{F}_c < F_t$, $prob > 5\%$)

$H_1 : Y_t$ n'est pas linéaire ($\hat{F}_c > F_t$, $prob < 5\%$)

Avec « $\hat{F}_c$ » : statistique calculée de Keenan.

Application du test sur notre processus ARCH(1)

► Estimation des coefficients pour AR(P) : les modèles sont :

Modèle 1 : $EA_t = \Phi_0 + \Phi_1 EA_{t-1} + \Phi_2 EA_{t-2} + \dots + \Phi_p EA_{t-p} + u_t$

Modèle 2 : $EA_t^2 = a_0 + a_1 EA_{t-1} + a_2 EA_{t-2} + \dots + a_p EA_{t-p} + e_t$

Sur EViews (nous n'utilisons pas Stata pour de raisons d'espace), faire :

$$\begin{cases} ls \ EA \ c \ EA(-1) \ EA(-2) \ EA(-3) \ EA(-4) \\ GENR \ EA2 = EA^2 \\ ls \ EA2 \ c \ EA(-1) \ EA(-2) \ EA(-3) \ EA(-4) \end{cases}$$

Notes :

- o à gauche (les résultats du modèle 1 estimé) et à droite (les résultats du modèle 2 estimé) ;
- o l'on vérifiera que le lag optimal (retard/décalage qui minimise le plus les coefficients d'Akaike et/ou Schwarz) est fixé à « 4 » pour les deux modèles.



Dependent Variable: EA Method: Least Squares Date: 01/01/14 Time: 21:31 Sample (adjusted): 7 200 Included observations: 194 after adjustments					Dependent Variable: EA2 Method: Least Squares Date: 01/01/14 Time: 21:34 Sample (adjusted): 7 200 Included observations: 194 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.	Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.071640	0.093761	-0.764074	0.4458	C	1.598597	0.197802	8.081816	0.0000
EA(-1)	0.021711	0.072779	0.298312	0.7658	EA(-1)	-0.401262	0.153537	-2.613450	0.0097
EA(-2)	0.037932	0.072831	0.520818	0.6031	EA(-2)	0.040156	0.153647	0.261354	0.7941
EA(-3)	-0.009434	0.072594	-0.129950	0.8967	EA(-3)	-0.045591	0.153146	-0.297695	0.7663
EA(-4)	0.020438	0.072638	0.281363	0.7787	EA(-4)	-0.071090	0.153240	-0.463911	0.6432
R-squared	0.002454	Mean dependent var	-0.077861		R-squared	0.036779	Mean dependent var	1.639747	
Adjusted R-squared	-0.018658	S.D. dependent var	1.281464		Adjusted R-squared	0.016393	S.D. dependent var	2.751173	
S.E. of regression	1.293363	Akaike info criterion	3.377804		S.E. of regression	2.728530	Akaike info criterion	4.870838	
Sum squared resid	316.1571	Schwarz criterion	3.462028		Sum squared resid	1407.081	Schwarz criterion	4.955061	
Log likelihood	-322.6470	F-statistic	0.116229		Log likelihood	-467.4713	F-statistic	1.804159	
Durbin-Watson stat	1.999227	Prob(F-statistic)	0.976652		Durbin-Watson stat	1.576230	Prob(F-statistic)	0.129705	

- o Les sommes des carrés des résidus/SSR (Sum squared resid) pour les deux modèles (t=5 à 200 observations) :

$$\text{SSR}(1)=316.1571 \text{ et } \text{SSR}(2)=1407.081$$

► Estimation du coefficient pour : $\hat{u}_t = \lambda \hat{e}_t + v_t$

_____ Commandes EViews :

$\left\{ \begin{array}{l} \text{GENR EA2=EA}^2 \\ \text{ls EA c EA(-1) EA(-2) EA(-3) EA(-4)} \\ \text{genr RESEA=RESID} \\ \text{ls EA2 c EA(-1) EA(-2) EA(-3) EA(-4)} \\ \text{genr RESEA2=RESID} \\ \text{ls RESEA RESEA2} \end{array} \right.$

Dependent Variable: RESEA Method: Least Squares Date: 01/01/14 Time: 22:02 Sample (adjusted): 7 200 Included observations: 194 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
RESEA2	-0.172427	0.031783	-5.425155	0.0000
R-squared	0.132320	Mean dependent var	4.81E-17	
Adjusted R-squared	0.132320	S.D. dependent var	1.279890	
S.E. of regression	1.192209	Akaike info criterion	3.194635	
Sum squared resid	274.3231	Schwarz criterion	3.211479	
Log likelihood	-308.8796	Durbin-Watson stat	2.071825	



► Calcul des paramètres et de la statistique « F » du test :

L'estimation précédente nous fournit : $\lambda = -0.172427$. Ce qui nous aide à calculer le coefficient « $\hat{\beta}^2$ » comme suit :

$$\hat{\beta}^2 = \hat{\lambda}^2 \left[\sum_{t=p+1}^T \hat{e}_t^2 \right] = (-0.172427)^2 \times (1407.081) = 41.8340$$

Et par endroit, la statistique « F » du test comme suit :

$$\hat{F} = \hat{\beta}^2 \left[\frac{T - 2p - 2}{\sum_{t=p+1}^T \hat{u}_t^2 - \hat{\beta}^2} \right] = \frac{190 \times 41.8340}{316.1571 - 41.8340} = 28.9748 \sim F_{\alpha(1,190)}$$

Les hypothèses du test sont :

H_0 : EA est linéaire ($\hat{F}_c < F_t$, $prob > 5\%$)

H_1 : EA n'est pas linéaire ($\hat{F}_c > F_t$, $prob < 5\%$)

_____ NB : au seuil de 5% : $F(1,190)=3,84$, et au seuil de 1% : $F(1,190)=6,64$.

_____ Décision : nous rejetons l'hypothèse nulle à 5% et 1% ($\hat{F}_c > F_t$) et affirmons que notre processus ARCH(1) « EA » n'est pas linéaire : ce qui confirme notre présomption sur Skewness.

Cas pratique 7.3 : Méthodologie de Box et Jenkins et Modélisation hétéroscédastique

(*Cfr page suivante*)



VII.1. METHODOLOGIE DE BOX ET JENKINS

VII.1.1. RAPPELS THEORIQUES

L'étude d'une série aux fins de prévision, par l'approche méthodologique de Box et Jenkins, passe par les cinq étapes qui suivent :

- o L'étude de la stationnarité (tests ADF ou PP ; graphiques) ;
- o L'identification du processus adéquat (lecture des corrélogrammes) ;
- o L'estimation du modèle/processus optimal retenu ;
- o L'inférence statistique (diagnostic/validation du modèle estimé) ;
- o La prévision.

A. Etude de la stationnarité

On se sert des tests formels et informels pour étudier la stationnarité de la série sous-étude (soit X_t). Si elle est non stationnaire, la stationnariser par les MCO ou l'écart à la tendance (TS), soit par la différenciation ou les filtres aux différences (DS).

B. Identification du processus adéquat (modèle optimal) et Estimation

Cette étape consiste à déterminer, dans la famille ARIMA⁽¹⁾, le processus générateur de X_t – par principe de parcimonie – ainsi que son ordre (décalage optimal). Le décalage optimal d'un ARIMA correspond au nombre (k) des coefficients d'autocorrélations simple ou partiel statistiquement significatifs (Cfr *corrélogramme de la série*). L'on notera que :

- (i) Pour un AR, l'ordre correspond au nombre des coefficients d'autocorrélation partiel (PAC) non nuls (différents de 0), tandis que ;
- (ii) Pour un MA, l'ordre correspond au nombre des coefficients d'autocorrélation simple (AC) non nuls (différents de 0) ;
- (iii) Pour un ARMA, les ordres sont respectivement définis suivant AR et MA. Toute fois, les processus mixtes (ARMA) présentent parfois des fonctions d'autocorrélation (AC et PAC) complexes/difficiles à interpréter, nécessitant une procédure itérative du genre identification-estimation-diagnostic.

Notons qu'en cas de décroissance trop lente d'une fonction/graphe d'autocorrélation, il est souhaitable de différencier la série avant d'identifier le modèle. En outre, en général, le nombre « k » de décalage optimal admissible – pour que les coefficients d'autocorrélation aient un sens – est donné par l'intervalle : $\frac{n}{6} \leq k \leq \frac{n}{3}$, soit $k = \frac{n}{5}$ ($\forall n \geq 150$) ; où n = nombre d'observations.

¹ Nous nous intéressons aux processus non saisonniers (où toute composante saisonnière a été éliminée). Les processus ARIMA saisonniers (SARIMA) font l'objet d'une spécification particulière, et ne sont pas abordés dans ce cadre.



Toute fois, la prudence s'impose quant à la l'identification⁽¹⁾. De manière générale, les fonctions d'autocorrélation doivent afficher des décroissances exponentielles monotones ou sinusoïdales. Cela doit s'observer comme suit :

- Pour un AR, la décroissance exponentielle doit s'observer sur la fonction AC ;
- Pour un MA, la décroissance exponentielle doit s'observer sur la fonction PAC (Partial autocorrelation).

Les conséquences d'une mauvaise identification sont entre autres l'invalidation des hypothèses de base des MCO (après estimation du modèle), et une prévision moins performante. Pour améliorer les résultats d'estimation et de prévision, le passage d'un processus à un autre se fait comme suit :

$$\begin{cases} \text{ARIMA}(p, d, q) \rightarrow \text{ARIMA}(p, d, 0); \\ \text{ARIMA}(p, d, q) \rightarrow \text{ARIMA}(p + 1, d - 1, 0) \end{cases}$$

Une fois le processus générateur de la série (X_t) identifié, passer à l'estimation par les MCO.

C. Diagnostic (inférence)

Pour confirmer le processus, il tient de tester la significativité des paramètres estimés (les paramètres statistiquement non significatifs doivent être éliminés/abandonnés). Aussi, tester :

- Si les résidus sont des bruits blancs⁽²⁾ gaussiens (homoscédasticité, absence d'autocorrélation des erreurs et normalité) ;
- La linéarité du modèle (test de significativité conjointe des paramètres) ;
- La stationnarité de la série dans le modèle estimé.

Si l'on considère le modèle ARIMA (2,0,0) estimé comme suit :

$$\hat{X}_t = \hat{\Phi}_1 X_{t-1} + \hat{\Phi}_2 X_{t-2} \dots \dots [7.1]$$

Tester la stationnarité de X_t à partir ce modèle revient à vérifier que les racines du polynôme retard associé à ce modèle sont à l'extérieur du cercle unité. C'est-à-dire :

$$(1 - \hat{\Phi}_1 L - \hat{\Phi}_2 L^2) = 0 \Rightarrow \begin{cases} |L_1| > 1 \\ |L_2| > 1 \end{cases} : \text{Condition de stationnarité de } X_t$$

$[\Phi(L) = 0 \text{ et que } |L_i| > 1]$

- Si la prévision dans l'échantillon est bonne. Voir si le coefficient de Theil est proche de 0 ou si le Mean Absolute Percent Error (MAPE) est faible.

En outre, l'on notera que les étapes d'identification, estimation et diagnostic peuvent être réitérées jusqu'à aboutir au résultat qui soit jugé satisfaisant (retenir le modèle adéquat).

¹ C'est l'étape la plus subjective dans la procédure de Box et Jenkins.

² Les résidus « u_t » sont des bruits blancs (c.à.d. processus strictement stationnaires) s'ils sont indépendants et identiquement distribués (iid) sur n'importe quel sous ensemble de temps (1,2,...,T).



D. Prédiction

Elle dépend de la nature de la série, selon qu'elle est stationnaire en tendance (TS) ou qu'elle est stationnaire en différence (DS).

(i) Si la série est un DS

Soit un processus ARIMA (1,1,0) stationnaire en différence, estimé comme suit :

$$\Delta \hat{X}_t = \hat{a}_0 + \hat{a}_1 \Delta X_{t-1} \dots \dots [7.2]$$

Sachant que $\Delta X_t = X_t - X_{t-1}$, [7.2] peut s'écrire alors :

$$\hat{X}_t = \hat{a}_0 + \hat{a}_1 \Delta X_{t-1} + \hat{X}_{t-1}$$

Ainsi :

- la prédiction à l'horizon « h » sera :

$$\hat{X}_{t+h} = \hat{a}_0 + \hat{a}_1 \Delta X_{t-1+h} + \hat{X}_{t-1+h}$$

- Si $h = 1$, alors :

$$\hat{X}_{t+1} = \hat{a}_0 + \hat{a}_1 \Delta X_t + \hat{X}_t$$

(ii) Si la série est un TS

Soit le processus ARIMA (1,0,0) stationnaire en tendance, estimé comme ci-dessous :

$$\hat{Z}_t = \hat{a}_1 Z_{t-1} \dots \dots [7.3]$$

avec : $\hat{Z}_t = X_t - \hat{X}_t \dots \dots [7.4a]$ et $\hat{X}_t = \hat{b}_0 + \hat{b}_1 t \dots \dots [7.5]$

[7.5] dans [7.4a] :

$$\hat{Z}_t = X_t - (\hat{b}_0 + \hat{b}_1 t) \dots \dots [7.4b]$$

De [7.4b], on déduit que :

$$X_t = \hat{Z}_t + \hat{b}_0 + \hat{b}_1 t \dots \dots [7.6], t = 0, 1, 2, \dots, T \text{ (tendance)}$$

Ainsi :

- A l'horizon « h », la prédiction de X_t sera :

$$X_{t+h} = \hat{Z}_{t+h} + \hat{b}_0 + \hat{b}_1 (t + h)$$

- Si $h = 2$ et que $T = 150$, alors :

$$\begin{cases} X_{t+1} = \hat{Z}_{t+1} + \hat{b}_0 + \hat{b}_1 (t + 1), \text{ avec } T + 1 = 151 \\ X_{t+2} = \hat{Z}_{t+2} + \hat{b}_0 + \hat{b}_1 (t + 2), \text{ avec } T + 2 = 152 \end{cases}$$

Comment trouver $\hat{Z}_{t+1}$ et $\hat{Z}_{t+2}$? Référons-nous à [7.3] et faisons :

$$\begin{cases} \hat{Z}_{t+1} = \hat{a}_1 Z_t \\ \hat{Z}_{t+2} = \hat{a}_1 \hat{Z}_{t+1} \end{cases}$$



NB : Si X_t est une série logarithmique (on écrira LX_t), dans ce cas : $X_{t+h} = e^{(LX_{t+h})}$.

La prévision (ponctuelle) une fois réalisée, il tient de visualiser graphiquement la série brute et la série prévue (ajustée) dans l'échantillon et en dehors de celui-ci (à l'horizon h)⁽¹⁾.

Il tient de préciser aussi qu'il est possible de construire un intervalle de prévision pour juger de la performance de la prévision effectuée. En fait, pour un ARIMA (2,0,0) [Cfr expression 7.1] estimé comme suit : $\hat{X}_t = \hat{\Phi}_1 X_{t-1} + \hat{\Phi}_2 X_{t-2}$, l'intervalle de prévision de X_t à l'horizon h (pour le seuil de confiance de 95%, soit $\alpha = 0,05$) est ⁽²⁾ :

$$\hat{X}_t(h) \pm 1,96\sigma_{\hat{e}} \left[\sum_{i=0}^{h-1} \gamma_i^2 \right]^{\frac{1}{2}}, \text{ avec } \begin{cases} \gamma_1 = \phi_1 \\ \gamma_2 = \phi_1^2 + \phi_2 \\ \gamma_3 = \phi_1^3 + 2\phi_1\phi_2 \end{cases} \text{ et } \sigma_{\hat{e}} = \frac{(SCR)}{T-k}$$

VII.1.2. ETUDE DE LA SERIE « EURUS » PAR L'APPROCHE METHODOLOGIQUE DE BOX ET JENKINS

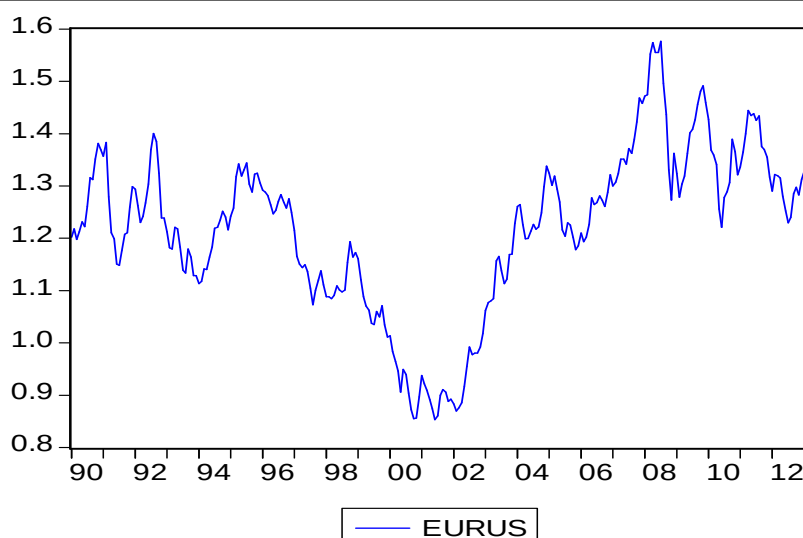
a) Etude de la stationnarité

► Evolution graphique de la parité Euro-dollar (série « EURUS ») : test informel

Sur EViews, taper :

```
{ create m 1990:01 2013:03
  data EURUS
  plot EURUS
```

Evolution mensuelle de la parité Euro-Usd (de janvier 1990 à mars 2013)



¹ On parle de la prévision in casting (dans l'échantillon) et out casting (hors échantillon).

² $\sigma_{\hat{e}}$ = l'écart-type de l'erreur de prévision; SCR = Somme des Carrés des Résidus (SCR); T et k = taille de l'échantillon et nombre des régresseurs.

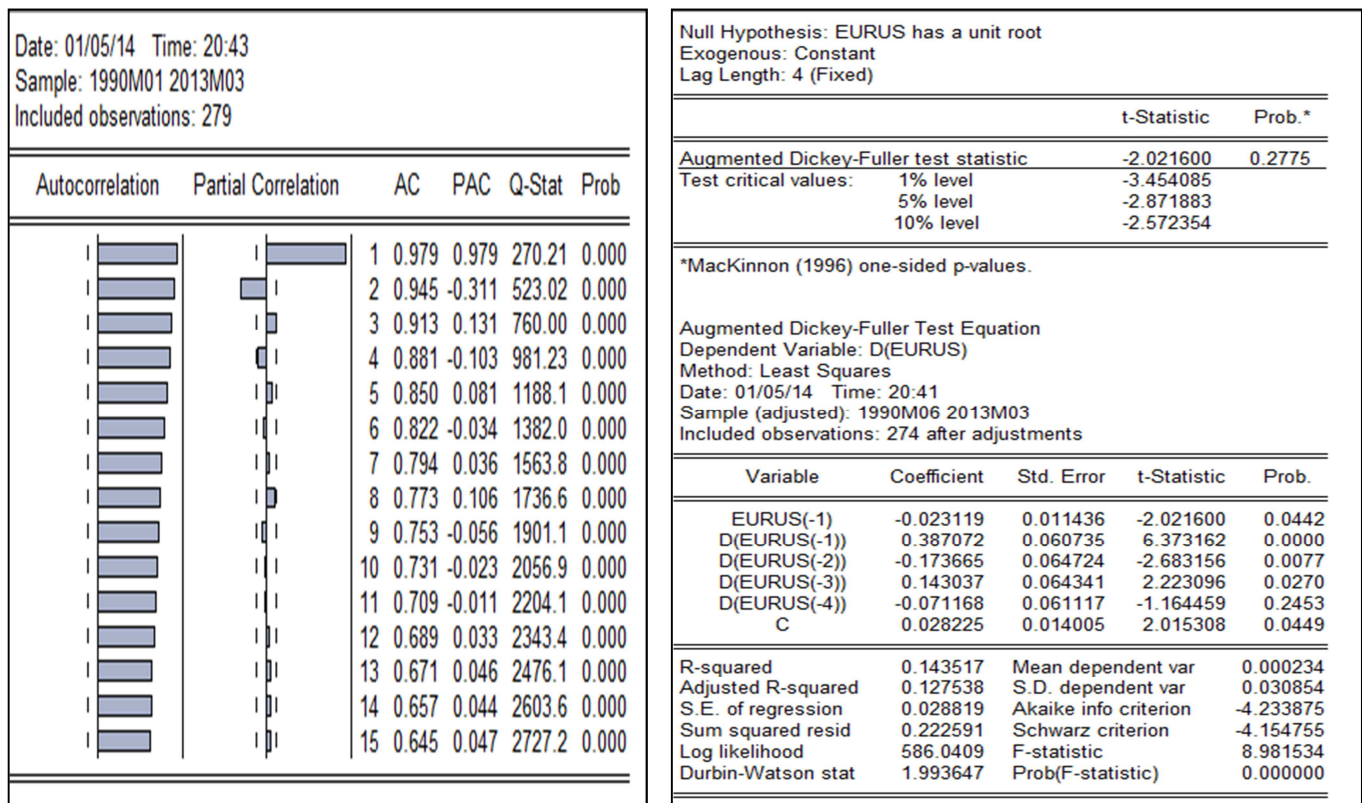


_____ Constat : à la lecture du graphique ci-haut, l'on présume une non stationnarité en moyenne (la série « EURUS » accuse une tendance évolutive/variable avec le temps) et en variance (à cause de la forte variabilité ou volatilité de la série). NB : l'échantillon a 279 observations.

► **Test d'Augmented Dickey-Fuller (ADF) : test formel**

_____ Sur Eviews, taper :

- { *ident EURUS* → (*lags* : 15) : pour obtenir la figure à gauche (corrélogramme);
uroot EURUS → (*user specified* : 4) : pour obtenir les résultats à droite :

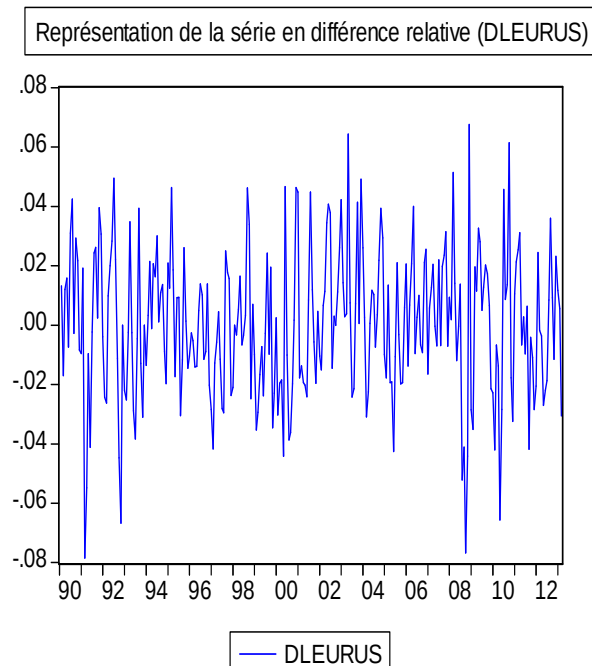


_____ Constat : le lag « 4 » dans « user specified » correspond au nombre de retard ou termes significativement différents de zéro (ceux qui traversent l'intervalle de confiance : Cfr corrélogramme partiel). Au regard des résultats issus du test « ADF » ($|ADF| < |Mackinnon|$ au seuil de 5%), l'on confirme que la série « EURUS » est non stationnaire en niveau du type DS (sans tendance, mais avec dérive). Pour la stationnariser, nous procédons par les filtres aux différences (différence première). Sur **EViews**, faire :

$$\begin{cases} \text{genr DLEURUS} = d(\log(\text{EURUS})) \\ \text{plot DLEURUS} \\ \text{uroot DLEURUS} \end{cases}$$

Ce qui nous donne le graphique (à gauche) et les résultats du test ADF (à droite) ci-dessous. Rien qu'à observer le graphique et, parce que $|ADF| > |Mackinnon|$ à 5%, l'on déduit que la série différenciée d'ordre 1 « DLEURUS » est stationnaire :





Null Hypothesis: DLEURUS has a unit root				
Exogenous: None				
Lag Length: 1 (Fixed)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-11.21635	0.0000
Test critical values:	1% level		-2.573367	
	5% level		-1.941978	
	10% level		-1.615931	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(DLEURUS)				
Method: Least Squares				
Date: 01/05/14 Time: 21:11				
Sample (adjusted): 1990M04 2013M03				
Included observations: 276 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLEURUS(-1)	-0.786490	0.070120	-11.21635	0.0000
D(DLEURUS(-1))	0.147063	0.059848	2.457270	0.0146
R-squared	0.356049	Mean dependent var	-4.93E-05	
Adjusted R-squared	0.353699	S.D. dependent var	0.029153	
S.E. of regression	0.023437	Akaike info criterion	-4.661783	
Sum squared resid	0.150506	Schwarz criterion	-4.635548	
Log likelihood	645.3261	Durbin-Watson stat	1.960152	

Constat : pour un lag fixé à « 1 » dans « user specified » (Cfr Corrélogramme ci-bas : identification), notre série différenciée est stationnaire sans trend, ni intercept (les modèles avec trend et/ou intercept se sont révélés non significatifs).

b) L'identification et l'estimation du processus adéquat/optimal

Pour identifier le processus adéquat dans la famille ARIMA – celui qui soit susceptible de nous reproduire le mode opératoire de la série « EURUS » – nous nous référons d'abord au corrélogramme de la série stationnaire « DLEURUS » (méthode quelque peu informelle), ensuite jugerons de la significativité/bonté du modèle ainsi identifié (au cas contraire, retenir le processus significatif par tâtonnement/réitération : Cfr rappels théoriques). A gauche (le corrélogramme de « DLEURUS » et à droite (les résultats du modèle identifié au départ : **arima (1,1,1)**) :

Sur **EViews**, faire : ident DLEURUS et ls DLEURUS ar(1) ma(1)

Date: 01/05/14 Time: 21:09					
Sample: 1990M01 2013M03					
Included observations: 278					
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.312	0.312	27.335	0.000	
2	-0.033	-0.144	27.634	0.000	
3	0.021	0.088	27.764	0.000	
4	-0.042	-0.094	28.255	0.000	
5	-0.065	-0.014	29.462	0.000	
6	-0.031	-0.018	29.736	0.000	
7	-0.081	-0.080	31.621	0.000	
8	-0.011	0.050	31.658	0.000	
9	0.042	0.010	32.171	0.000	
10	0.007	-0.005	32.186	0.000	
11	-0.046	-0.056	32.793	0.001	
12	-0.027	-0.003	33.009	0.001	
13	-0.046	-0.048	33.629	0.001	
14	-0.090	-0.070	36.025	0.001	
15	0.011	0.067	36.060	0.002	

Dependent Variable: DLEURUS				
Method: Least Squares				
Date: 01/05/14 Time: 21:35				
Sample (adjusted): 1990M03 2013M03				
Included observations: 277 after adjustments				
Convergence achieved after 10 iterations				
Backcast: 1990M02				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.132893	0.145611	-0.912653	0.3622
MA(1)	0.512801	0.126448	4.055441	0.0001
R-squared	0.128512	Mean dependent var	0.000223	
Adjusted R-squared	0.125343	S.D. dependent var	0.024883	
S.E. of regression	0.023272	Akaike info criterion	-4.675959	
Sum squared resid	0.148933	Schwarz criterion	-4.649793	
Log likelihood	649.6203	Durbin-Watson stat	1.989823	
Inverted AR Roots	-.13			
Inverted MA Roots	-.51			



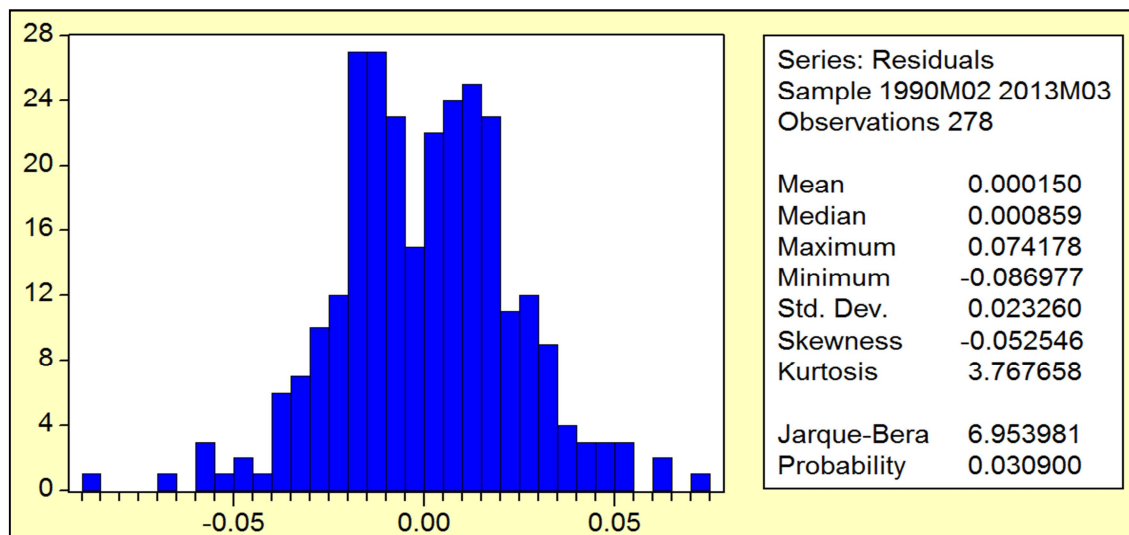
_____ Constat : notre « *arima(1,1,1)* » nous donne des résultats non satisfaisant pour la composante autorégressive qui est statistiquement non significative. Tout compte fait, nous avons plutôt retenu le processus « *arima(0,1,1)* » comme adéquat. Sur **EViews**, faire : *ls DLEURUS ma(1)* :

Dependent Variable: DLEURUS				
Method: Least Squares				
Date: 01/05/14 Time: 21:42				
Sample (adjusted): 1990M02 2013M03				
Included observations: 278 after adjustments				
Convergence achieved after 7 iterations				
Backcast: 1990M01				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA(1)	0.402891	0.055171	7.302550	0.0000
R-squared	0.123917	Mean dependent var		0.000270
Adjusted R-squared	0.123917	S.D. dependent var		0.024851
S.E. of regression	0.023260	Akaike info criterion		-4.680552
Sum squared resid	0.149868	Schwarz criterion		-4.667503
Log likelihood	651.5967	Durbin-Watson stat		2.042118
Inverted MA Roots	-.40			

c) L'inférence statistique (diagnostic/validation du modèle estimé)

Nous faisons seulement le test de normalité de Jarque-Bera (le test de bruit blanc et celui de présence d'effets ARCH étant concluant).

_____ Test de normalité : Sur **Eviews**, dans l'output des résultats, suivre : View/Residual Tests/Histogram-Normality Test :



_____ Constat : les résidus du modèle optimal estimé ne sont pas normalement distribués (présomption d'une non linéarité) : ils sont leptokurtiques ($Skewness > 3$) et légèrement étalé vers la droite ($Kurtosis \neq 0$ et *negative* : dissymétrique à gauche).

D'où, nous rejetons la spécification ARIMA au profit de la modélisation hétéroscédastique (ARCH) – qui est adaptée à l'étude des séries chronologiques



accusant une forte variabilité/volatilité (impliquant la non stationnarité, la non normalité ou la non linéarité) – pour prévoir l'évolution future de la parité Euro-dollar (soit la série « EURUS »).

VII.2. MODELISATION HETEROSCEDASTIQUE

VII.2.1. RAPPELS THEORIQUES

A. Notes

L'étude des séries chronologiques a conduit au développement de plusieurs modèles parmi lesquels la spécification autorégressive – la modélisation ARMA en général – a été largement utilisée, à cause notamment de ses propriétés statistiques⁽¹⁾ qui rendent facile l'estimation, l'inférence statistique et la prévision (les séries et résidus étant générés par des processus stationnaires). Malgré ces avantages, les modèles ARMA(p,q) souffrent de la non prise en compte de certaines contraintes structurelles liées au phénomène faisant l'objet de la modélisation. Ces contraintes peuvent traduire le caractère volatile de certaines variables ou même le comportement rationnel des agents économiques, et impliquent parfois le recours à des modèles non linéaires susceptibles de rendre inadéquat la spécification ARMA. Dans ce contexte, comme le propose Engle (1982) – pour intégrer la volatilité dans la spécification ARMA – il tient de recourir au modèle Autorégressif Conditionnellement Hétéroscédastique (ARCH). En effet, ce type des modèles est adapté aux chroniques présentant les deux caractéristiques qui suivent :

- ▶ La non stationnarité (surtout en variance) ;
- ▶ La non normalité de la distribution d'une variable/série, à la suite de son caractère leptocurtique, pouvant conduire à la non linéarité de la variable/série sous étude.

Précisons que ces caractéristiques sont propres aux données financières et monétaires (inflation, taux de change, cours boursiers,...) dont l'évolution est souvent non linéaire et volatile. Il s'en suit que, pour ce type des données, les MCO ne sont pas valides à cause de l'hétéroscédasticité et la non linéarité dans les variables et certains paramètres. D'où, le recours à la méthode du maximum de vraisemblance tient lieu.

Par ailleurs, il est à préciser que le modèle ARCH consiste à représenter, de façon autorégressive, la variance de l'erreur étant donné la connaissance de son information passée⁽²⁾. Dans le groupe/famille de modèles ARCH, il est à distinguer les modèles ARCH linéaires (on y trouve les modèles : ARCH(q), ARCH-M, GARCH-M,

¹ A cause de la linéarité dans les variables et dans certains paramètres, les MCO sont facilement applicables).

² On parle mieux de la régression ou représentation autorégressive de la variance des erreurs conditionnellement à son information passée.



GARCH(p,q) et IGARCH(p,q)) des modèles ARCH non linéaires (on y trouve les modèles : TARARCH(p,q), TGARCH(p,q) et EGARCH(p,q))⁽¹⁾.

En fait, les ARCH linéaires spécifient la variance conditionnelle des erreurs sous la forme quadratique et ne prend en compte que l'ampleur des valeurs passées ; tandis que, les ARCH non linéaires tiennent à la fois compte du signe et de l'ampleur des valeurs passées en levant l'hypothèse des spécifications quadratiques symétriques (on fait des spécifications asymétriques des erreurs).

B. Modélisation

B.1. Les processus ARCH linéaires

L'hypothèse fondamentale sous-tendant les ARCH linéaires est la symétrie des spécifications quadratiques de la variance conditionnelles des erreurs.

1) Modèle ARCH(q)

Rappelons qu'un modèle de type ARCH(q) consiste à spécifier la variance des erreurs de façon autorégressive conditionnellement à son information passée. Une telle spécification peut généralement s'écrire :

$$\text{var}(e_t/I_{t-1}) = h_t^2 = a_0 + a_1 e_{t-1}^2 + a_2 e_{t-2}^2 + \dots + a_q e_{t-q}^2 \dots \dots [1]$$

Où $\text{var}(e_t/I_{t-1}) = 0$; $a_0 > 0$ et $a_i \geq 0$ ($i = 1 \dots q$),

$$\sum_{i=1}^q a_i < 1 : \text{condition de stationnarité}$$

Avec « e_t ($t \in \mathbb{Z}$) », le résidu d'un modèle de régression général : $Y_t = AX_t + e_t$ ou d'un ARMA(p,q) :

$$Y_t = \lambda + \sum_{i=1}^p \phi_i Y_{t-i} + \sum_{i=1}^q \theta_i e_{t-i} + e_t \dots \dots [1]'$$

2) Modèle ARCH Generalised (GARCH(p,q))

Il s'agit d'un modèle ARCH généralisé ; car, dans ce type des modèles, l'information – plus éloignée dans le passé – sur la variance conditionnelle des erreurs est prise en compte dans la spécification de celle-ci en y incluant les valeurs des variances décalées. Selon BOLLERSLEV (1986), le modèle GARCH(p,q) s'écrit :

$$h_t^2 = a_0 + \sum_{i=1}^q a_i e_{t-i}^2 + \sum_{i=1}^p b_i h_{t-i}^2 \dots \dots [2]$$

¹ Signalons de passage qu'il existe des modèles : Threshold Autoregressive/Autorégressif à seuils (TAR), Smooth Transition Autoregressive/Autorégressif à transition douce (STAR), Logistic STAR/STAR Logistique (LSTAR), Exponential STAR/STAR exponentiel (ESTAR).



Où :

$$a(L) + b(L) = \sum_{i=1}^q a_i L^i + \sum_{i=1}^p b_i L^i, \text{ et}$$

$$a(1) + b(1) = \sum_{i=1}^q a_i + \sum_{i=1}^p b_i < 1 \text{ (condition de stationnarité)}$$

_____ NB : $L^i e_t = e_{t-i}$ ($i \geq 0$).

3) **Modèle GARCH integrated (IGARCH(p,q))**

Le modèle GARCH intégré ou IGARCH est une spécification GARCH pour des processus non stationnaires en niveau (Engle et Bollerslev, 1986). L'on suppose donc que « $\sum_{i=1}^q a_i + \sum_{i=1}^p b_i > 1$ », à tel enseigne qu'un choc sur « h_t^2 » se répercute sur les valeurs de h_{t+m}^2 (m : horizon de prévision) de façon explosive, sans s'estomper dans le temps. Ce phénomène est appelé « persistance ». Il va sans dire que dans le modèle IGARCH, l'effet de persistance dans la variance des erreurs – caractéristiques des processus non stationnaires – est pris en compte.

En effet, les variances conditionnelles prévues jusqu'à l'horizon « m » s'expriment comme suit :

$$E(h_{t+m}^2 / I_t) = (a + b)^m h_t^2 + a_0 \left[\sum_{i=0}^{m-1} (a + b)^i \right] \dots \dots \dots [3]$$

Si l'on considère un modèle IGARCH(1,1), où « $a + b = 1$ », alors :

$$E(h_{t+m}^2 / I_t) = h_t^2 + a_0 m \dots \dots \dots [3]'$$

Avec :

$$h_t^2 = a_0 + a_1 e_{t-1}^2 + b_1 h_{t-1}^2 \dots \dots \dots [3]''$$

4) **Modèle ARCH in Mean (ARCH-M) et GARCH in Mean (GARCH-M)**

Les modèles ARCH et GARCH avec effet de moyenne sont de spécifications dans lesquelles les effets ARCH et GARCH respectivement influencent aussi la moyenne conditionnelle (Engle, Lilien et Robins, 1987).

Considérons la régression suivante :

$$Y_t = CX_t + f(h_t^2) + e_t \dots \dots \dots [4]$$

Où : $Y_t / I_{t-1} \rightarrow N(u_t ; h_t^2)$, avec : $u_t = E(Y_t) = CX_t + f(h_t^2)$ et $f(h_t^2)$: fonction de « h_t^2 ».

Dans cette régression, l'espérance conditionnelle « u_t » est exprimée en fonction de la variance conditionnelle « h_t^2 ». Autrement dit, le niveau atteint par la variable est expliquée par sa forte variabilité (volatilité). Considérant l'expression (4) : (i) on fera allusion à un modèle de type ARCH-M si la variance conditionnelle des erreurs « h_t^2 » – tenant compte des effets ARCH – peut s'écrire : $h_t^2 = a_0 + a_1 e_{t-1}^2 + a_2 e_{t-2}^2 + \dots + a_q e_{t-q}^2$, et (ii) on fera allusion à un modèle de type GARCH-M si la



variance conditionnelle des erreurs « h_t^2 » – tenant compte des effets GARCH – peut s'écrire :

$$h_t^2 = a_0 + \sum_{i=1}^q a_i e_{t-i}^2 + \sum_{i=1}^p b_i h_{t-i}^2$$

B.2. Les processus ARCH non linéaires

_____ L'hypothèse à la base des ARCH non linéaires est la prise en compte de l'asymétrie de l'information ou effet de levier dans les spécifications quadratiques de la variance conditionnelle des erreurs.

5) Modèles Exponential GARCH (EGARCH)

Le modèle GARCH exponentiel, difficile à manier ou à interpréter, est une spécification adaptée au modèle GARCH où « a_i et b_i » sont négatifs, levant ainsi les contraintes de non négativité imposées aux paramètres. Ce type des modèles s'expriment comme suit :

$$\log(h_t^2) = a_0 + \sum_{i=1}^q a_i \{\gamma V_{t-i} + \lambda[|V_{t-i}| + E|V_{t-i}|]\} + \sum_{i=1}^p b_i \log(h_{t-i}^2) \dots \dots [5]$$

Avec :

$$V_{t-i} = \frac{e_{t-i}}{h_{t-i}} : \text{l'erreur standardisée}$$

6) Modèles Threshold ARCH (TARCH) et Threshold GARCH (TGARCH)

La modélisation ARCH ou GARCH à seuils consiste à intégrer l'effet d'asymétrie dans les spécifications quadratiques de la variance conditionnelle des erreurs, si bien que le signe et l'amplitude d'un choc dans les erreurs décalées soient déterminants quant à ses effets sur la variance conditionnelle au temps t .

_____ Le modèle ARCH à seuils (TARCH(q)) s'écrit :

$$h_t^2 = a_0 + a_1 e_{t-1}^2 + \lambda e_{t-1}^2 D_{t-1} \dots \dots [6]$$

Où $D_{t-1} = \begin{cases} 1, & \text{si } e_{t-1} < 0 \\ 0, & \text{ailleurs} \end{cases}$: variable dichotomique permettant de capter l'effet de levier ou d'asymétrie.

_____ Le modèle GARCH à seuils (TGARCH(p, q)) s'écrit :

$$h_t^2 = a_0 + \sum_{i=1}^q a_i e_{t-i}^2 + \lambda e_{t-1}^2 D_{t-1} + \sum_{i=1}^p b_i h_{t-i}^2$$



C. Arbitrage entre la modélisation ARMA sans effets ARCH et celle avec effets ARCH

Lorsqu'on cherche à déterminer, dans la famille ARIMA, le processus qui permet de reproduire au mieux le mode opératoire d'une chronique (par principe de parcimonie), il est indispensable d'étudier la volatilité de celle-ci ou d'y tester la présence d'effets ARCH, surtout s'il s'agit d'une variable financière ou monétaire. En fait, la présence d'effets ARCH est synonyme d'autocorrélation de la variance des résidus, avec comme corollaire l'inflation de la variance ou la non stationnarité en variance, la non normalité, et donc la non linéarité de la série concernée. Alors, pour tester la présence d'effets ARCH dans un processus, l'on peut procéder par une série de tests dont :

- L'analyse graphique des séries brutes et stationnaires ;
- L'étude des statistiques descriptives de la série ;
- Les tests de marche aléatoire et de présence d'effets ARCH d'ordre supérieur à 3 ; et,
- La spécification autorégressive de la série filtrée (stationnaire) au carré.

(i) L'analyse graphique

En représentant sur un même graphique les séries brute et filtrée, l'on aura à présumer l'existence d'une hétéroscédasticité conditionnelle si la série laisse présager des fortes variabilités ou une non stationnarité en variance.

(ii) L'étude des statistiques descriptives

L'une des caractéristiques des processus ARCH est la non normalité (ou non linéarité) de la série. La statistique de Jarque-Bera, ainsi que sa probabilité associée conduisent l'inférence.

(iii) Le test de marche aléatoire

Basé sur la statistique de Ljung-Box, le test de bruit blanc permet de juger de l'hétéroscédasticité de la variance conditionnelle des erreurs lorsque l'on s'intéresse aux corrélogrammes des carrés des résidus. Ces derniers permettent de tester :

H₀ : la spécification est du type ARMA (termes du corrélogramme significativement nuls : $prob > 5\%$) ;

H₁ : la spécification est du type ARCH (termes du corrélogramme significativement différent de zéro : $prob < 5\%$).

Aussi, le test ARCH d'hétéroscédasticité d'ordre > 3 – appelé aussi test du multiplicateur de vraisemblance – renseigne sur la nécessité ou pas d'une modélisation du type ARCH. Les hypothèses du test sont :

H₀ : Absence d'effets ARCH d'ordre > 3 (4) : $prob > 5\%$, $F_c < F_t$) ;

H₁ : Existence d'effets ARCH d'ordre > 3 (4) : $prob < 5\%$, $F_c > F_t$).

En outre, si l'on considère la spécification des erreurs du type ARCH(q) suivante :



$$h_t^2 = a_0 + a_1 e_{t-1}^2 + a_2 e_{t-2}^2 + \dots + a_q e_{t-q}^2$$

Les hypothèses à vérifier sont :

$H_0 : a_1 = a_2 = \dots = a_q = 0$: Modélisation ARMA des erreurs ($prob > 5\%$, $F_c < F_t$) ;

$H_1 : a_1 \neq a_2 \neq \dots \neq a_q \neq 0$: Modélisation ARCH des erreurs ($prob < 5\%$, $F_c > F_t$).

Il s'agit là d'un test de Fisher. L'on peut aussi recourir au multiplicateur de Lagrange (LM) – qui suit la statistique de χ^2 à q degré de liberté – pour vérifier si :

H_0 : Modèle linéaire ARMA ($LM < \chi_q^2$) ;

H_1 : Modèle linéaire ARCH ($LM > \chi_q^2$).

(iv) Spécification autorégressive de la série filtrée au carré

Considérons un AR(1) tel que : $Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t$. La spécification autorégressive d'ordre 1 de la série « Y_t » filtrée au carré s'écrit : $Z_t = \phi_0 + \phi_1 Z_{t-1}$, avec : $Z_t = (\Delta Y_t)^2$. Les hypothèses sont :

$H_0 : \hat{\phi}_1 = 0$: Absence d'hétéroscédasticité conditionnelle (Modélisation ARMA sans effets ARCH ($prob > 5\%$, $|t_c| < |t_t|$) ;

$H_1 : \hat{\phi}_1 \neq 0$: Existence d'hétéroscédasticité conditionnelle (Modélisation ARMA avec effets ARCH ($prob < 5\%$, $|t_c| > |t_t|$).

D. Choix ou sélection du processus ARCH adéquat (optimal)

En présence d'une hétéroscédasticité conditionnelle de la variance des résidus, le choix sur un processus ARCH adéquat est guidé par plusieurs critères, notamment le caractère marche aléatoire des résidus issus de la variance conditionnelle, les meilleurs propriétés prédictives (MAPE faible et l'indice d'inégalité de Theil proche de zéro) et la meilleure qualité de l'équation de la moyenne (faible AIC).

Par ailleurs, l'arbitrage peut aussi se faire comme suit :

► **Modèle ARCH vs GARCH**

Partant de la spécification des erreurs de la forme GARCH(p,q) suivante :

$$h_t^2 = a_0 + \sum_{i=1}^q a_i e_{t-i}^2 + \sum_{i=1}^p b_i h_{t-i}^2 \dots \dots \dots [Cfr expression 2]$$

L'on peut tester si :

$H_0 : b_i = 0$: Modélisation ARCH des erreurs ($LM < \chi^2(p)$, $prob > 5\%$) ;

$H_1 : b_i \neq 0$: Modélisation GARCH des erreurs ($LM > \chi^2(p)$, $prob < 5\%$).

► **Modèle GARCH vs IGARCH**

Si la série brute sous-étude est non stationnaire, le modèle GARCH fait place au modèle GARCH intégré (IGARCH).



► **Modèle GARCH vs EGARCH**

Le modèle GARCH fait place au modèle EGARCH si l'effet de levier est validé : c.à.d. « γ » est statistiquement significatif (Cfr expression 5).

► **Modèle GARCH vs TGARCH**

Le modèle GARCH fait place au modèle TGARCH en présence d'effets seuils, c.à.d. quand « λ » est statistiquement significatif dans « l'expression 6 ».

► **Modèle ARCH ou GARCH vs ARCH-M ou GARCH-M**

Les modèles ARCH ou GARCH font place aux modèles ARCH-M ou GARCH-M si le paramètre associé à la variance conditionnelle « h_t^2 » est statistiquement significatif (Cfr expression 4), ou si graphiquement les fortes inflations lisibles sur la série brute correspondent aux fortes variabilités (lisibles sur la série filtrée) de la série concernée.

E. Estimation des modèles ARCH⁽¹⁾

Pour estimer les paramètres des modèles ARCH, l'on recourt souvent à la méthode de moindres carrés pondérés ou à des algorithmes fondés sur la méthode du maximum de vraisemblance.

F. Prévision des modèles ARCH

La prévision à travers une modélisation ARCH sera dite meilleure si elle minimise le MAPE (*Mean Absolute Pourcentage Error*) et présente un coefficient de Theil proche de zéro : faut-il encore que les résidus de la variance conditionnelle des erreurs soient des bruits blancs (gaussiens).

VII.2.2 ETUDE DE LA SERIE BRUTE « EURUS » PAR LA MODELISATION HETEROSCEDASTIQUE (ARCH)

Rappelons que, faisant partie de la famille ARIMA, la modélisation ARCH est de nature à éliminer l'hétéroscédasticité conditionnelle. Aussi, nous y recourons pour deux raisons :

- o Du point de vue « statistique » : Le caractère non normal/non linéaire de notre série « EURUS » ;
- o Du point de vue théorique : La paire « Euro-dollars » étant la plus traidée (une bonne part des transactions sur les marchés financiers/Forex⁽²⁾ sont axées sur la parité « EUR/US » : soit 30%), elle est victime des fortes spéculations et, par endroit, d'une forte volatilité (ce qui rend inadéquat l'approche de Box et Jenkins pour la modélisation de ce produit du Forex : la paire « EUR/US »).

¹ GARCH est identifié suivant la même logique que le processus ARMA.

² Selon « FxPro » (plateforme d'information financière) ou consulter les analyses des Economistes sur « fr.investing.com/economic-calendar » (A savoir : environ 100 paires des devises font l'objet de 95% des spéculations sur le Forex).



Précisons aussi que nous allons estimer six modèles (ARCH, GARCH, TGARCH, EGARCH, ARCH-M et GARCH-M) parmi lesquels nous retiendrons le plus performant, et allons procéder comme suit :

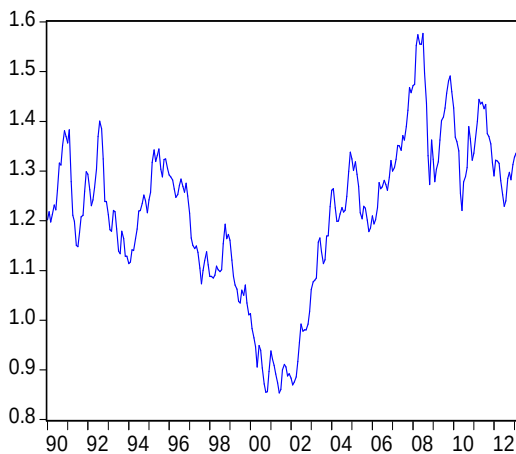
- (i) Représentations graphiques des séries « EURUS » et « DEURUS » ;
- (ii) Statistiques descriptives et test de normalité de la série « EURUS » ;
- (iii) Etude de la volatilité de la série « EURUS » ;
- (iv) Estimation des modèles (6) et recherche du modèle optimal
- (v) Prévision

A. Représentations graphiques des séries « EURUS » et « DEURUS »

Sur **EViews**, faire :

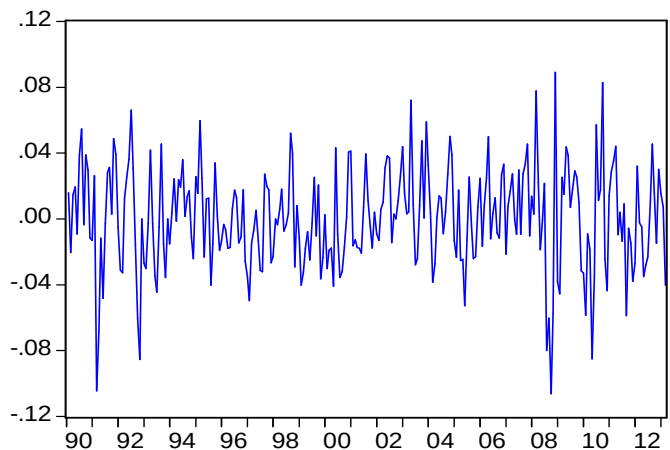
$\left\{ \begin{array}{l} \text{genr DEURUS} = d(\text{EURUS}) \\ \text{plot DEURUS EURUS} \text{ (« options » pour les axes : Normalized data)} \end{array} \right.$

Evolution mensuelle de la parité Euro-Usd (de janvier 1990 à mars 2013)



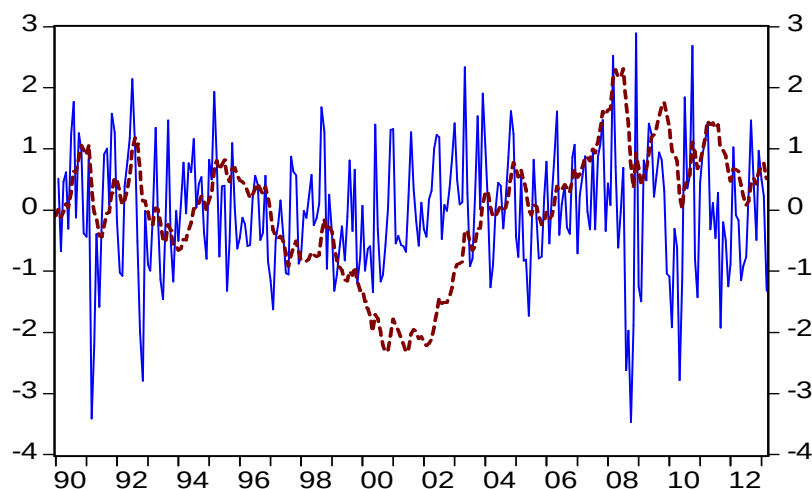
— EURUS

Représentation de la série en différence absolue (DEURUS)



— DEURUS

Evolution comparée de la série brute (EURUS) et celle en différence absolue (DEURUS)

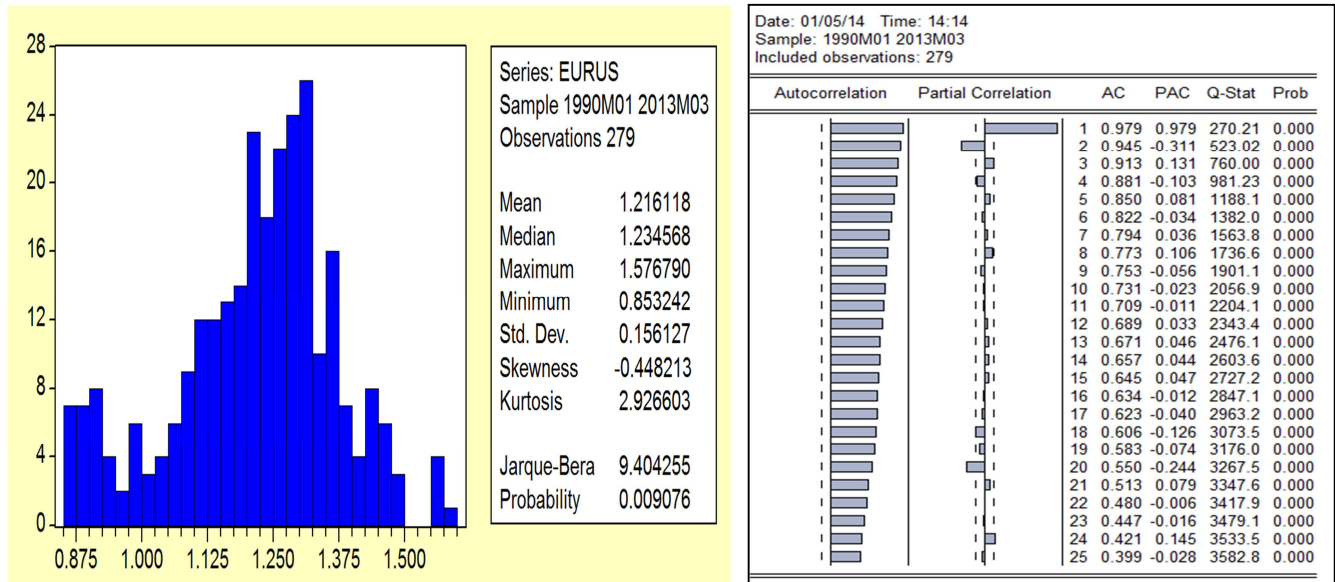


— DEURUS — EURUS



B. Statistiques descriptives et test de normalité de la série « EURUS »

Sur **EViews**, dans l'output des données, cliquer (pour obtenir la figure à gauche) : View/Descriptive Statistics/Histogram and stats ; et cliquer (pour obtenir la figure à droite) : View/Correlogram... → (level ; lag=25) :



Constats :

- o (i) Skewness (mesure d'asymétrie, il est nul pour une distribution normale) = -0.448213 : c.à.d. distribution dissymétrique à gauche (étalée à droite). D'où, présomption de non linéarité (Skewness non nul) ; (ii) Kurtosis (mesure d'applatissage, il avoisine « 3 » pour une distribution normale) = 2.926603 $\approx$ (<) 3 : c.à.d. distribution platykurtique (centre moins élevé) ; et (iii) Jarque-Bera (test de normalité, avec H_0 : le processus est normalement distribué). La probabilité associée est de 0.009076 < 5% : rejet de l'hypothèse de normalité pour la série « EURUS » sous-étude ;
- o Notre série brute « EURUS » suivrait un processus du type « AR(2) » (Cfr correlogramme). Elle accuse aussi une non stationnarité (décroissance lente des termes du correlogramme simple), serait-il beaucoup plus en variance.

C. Etude de la volatilité de la série « EURUS »

Pour étudier la volatilité de notre série « EURUS », régressons « DEURUS » au carré sur cette même série décalée (cette façon de procéder permet de juger de l'autocorrélation de la variance des résidus/hétéroscédasticité).

Sur **EViews**, faire :

$$\begin{cases} \text{genr DEURUS2=DEURUS}^2 \\ \text{ls DEURUS c DEURUS2(-1)} \end{cases}$$



Dependent Variable: DEURUS2 Method: Least Squares Date: 01/05/14 Time: 14:07 Sample (adjusted): 1990M03 2013M03 Included observations: 277 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000769	0.000107	7.176047	0.0000
DEURUS2(-1)	0.184564	0.059266	3.114178	0.0020
R-squared	0.034065	Mean dependent var		0.000942
Adjusted R-squared	0.030552	S.D. dependent var		0.001549
S.E. of regression	0.001525	Akaike info criterion		-10.12609
Sum squared resid	0.000640	Schwarz criterion		-10.09992
Log likelihood	1404.463	F-statistic		9.698106
Durbin-Watson stat	2.058548	Prob(F-statistic)		0.002039

_____ **Constat :** le paramètre associé à « DEURUS2(-1) » est statistiquement significatif, ce qui permet d'accepter l'hypothèse d'hétéroscédasticité conditionnelle (la variation de « EURUS » au temps « t » est fonction de son évolution au temps « t-1 »).

D. Estimation des modèles et recherche du modèle optimal

1) Estimation des modèles

a) Estimation du modèle ARCH

Nous estimons un AR(1)⁽¹⁾ ou un AR(2) à partir duquel nous effectuons le test de présence d'effets ARCH pour juger de la nécessité d'une modélisation ARCH.

► Estimation du modèle AR(1) :

_____ Sur **EViews**, faire : *ls EURUS c EURUS(-1)*

Dependent Variable: EURUS Method: Least Squares Date: 01/05/14 Time: 14:17 Sample (adjusted): 1990M02 2013M03 Included observations: 278 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.023235	0.014424	1.610806	0.1084
EURUS(-1)	0.981167	0.011767	83.38097	0.0000
R-squared	0.961817	Mean dependent var		1.216169
Adjusted R-squared	0.961679	S.D. dependent var		0.156406
S.E. of regression	0.030618	Akaike info criterion		-4.127305
Sum squared resid	0.258735	Schwarz criterion		-4.101207
Log likelihood	575.6954	F-statistic		6952.386
Durbin-Watson stat	1.361607	Prob(F-statistic)		0.000000

¹ Le processus générateur de la série « EURUS » était un **arima(1,1,1)** si l'on s'en tient aux résultats issus de la procédure Box-Jenkins appliquée à cette série. C'est ainsi que nous estimons les modèles AR(1) et AR(2), le processus arima(1,1,1) étant non significatif.



► Estimation du modèle AR(2) :

Sur **EViews**, faire : $ls\ EURUS\ c\ EURUS(-1)\ EURUS(-2)$

Dependent Variable: EURUS Method: Least Squares Date: 01/16/14 Time: 14:01 Sample (adjusted): 1990M03 2013M03 Included observations: 277 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.030959	0.013764	2.249333	0.0253
EURUS(-1)	1.299107	0.057303	22.67079	0.0000
EURUS(-2)	-0.324468	0.057363	-5.656356	0.0000
R-squared	0.965841	Mean dependent var	1.216163	
Adjusted R-squared	0.965592	S.D. dependent var	0.156689	
S.E. of regression	0.029065	Akaike info criterion	-4.227801	
Sum squared resid	0.231466	Schwarz criterion	-4.188551	
Log likelihood	588.5504	F-statistic	3873.704	
Durbin-Watson stat	1.910455	Prob(F-statistic)	0.000000	

Commentaires : le modèle AR(2) est plus significatif que le modèle AR(1), ce qui nous amène à retenir le processus AR(2) qui est optimal.

► Inférences sur le modèle AR(2) estimé :

Sur **EViews**, dans l'output de l'estimation (figure à gauche : obtenir le corrélogramme des résidus (simple)), suivre : *View/Residual Tests/Correlogram-Q-statistics* → *lags=15*. Aussi, suivre (figure à droite : obtenir le corrélogramme des résidus aux carrés) : *View/Residual Tests/CorrelogramSquared Residuals* → *lags=15* :

Date: 01/16/14 Time: 23:22						
Sample: 1990M03 2013M03						
Included observations: 277						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.041	0.041	0.4669	0.494
		2	-0.140	-0.142	5.9931	0.050
		3	0.089	0.104	8.2235	0.042
		4	-0.006	-0.037	8.2328	0.083
		5	-0.053	-0.024	9.0353	0.108
		6	0.026	0.016	9.2291	0.161
		7	-0.098	-0.112	12.002	0.101
		8	-0.005	0.024	12.008	0.151
		9	0.064	0.028	13.199	0.154
		10	0.009	0.024	13.221	0.212
		11	-0.047	-0.039	13.868	0.240
		12	0.016	0.007	13.941	0.305
		13	0.015	0.006	14.006	0.373
		14	-0.079	-0.081	15.857	0.322
		15	0.037	0.052	16.260	0.365

Date: 01/16/14 Time: 23:24						
Sample: 1990M03 2013M03						
Included observations: 277						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.034	0.034	0.3148	0.575
		2	0.176	0.175	9.0477	0.011
		3	0.097	0.089	11.702	0.008
		4	0.065	0.031	12.890	0.012
		5	0.227	0.202	27.574	0.000
		6	-0.055	-0.090	28.447	0.000
		7	0.061	-0.015	29.507	0.000
		8	0.021	0.009	29.638	0.000
		9	0.096	0.085	32.297	0.000
		10	-0.026	-0.082	32.486	0.000
		11	0.010	0.015	32.517	0.001
		12	-0.033	-0.050	32.832	0.001
		13	-0.043	-0.053	33.383	0.001
		14	0.040	0.025	33.850	0.002
		15	-0.036	0.021	34.227	0.003



Note : La statistique Q de Ljung-Box indique deux termes statistiquement différents de zéro (autant pour le corrélogramme des résidus que pour celui des résidus aux carrés); *Ce qui amène à présumer la présence d'effets ARCH d'ordre 2* dans notre série « EURUS » (le test ARCH ci-dessous nous en dit plus).

Test de présence d'effets ARCH sur le modèle AR(2) estimé : Sur EViews, dans l'output des résultats, suivre : View/Residuals Tests/ARCH LM Test... → lag(1) :

Note : la probabilité critique étant > 5% pour un ARCH(1), nous rejetons l'hypothèse de présence d'effets ARCH d'ordre 1 et testons cette hypothèse pour un ordre > 1. Pour un ARCH(2), le coefficient associé à « RESID²(-1) » est statistiquement significatif au seuil de 5%, ce qui nous amène à tester la présence d'effets ARCH dans notre processus « EURUS » jusqu'à l'ordre 4. Pour ce décalage (4), seul le coefficient associé à « RESID²(-1) » s'est avéré statistiquement significatif au seuil de 5% : d'où, nous confirmons que notre série brute « EURUS » suit un processus ARCH d'ordre 2.

ARCH (1) Test:				
F-statistic	0.289929	Probability	0.590703	
Obs*R-squared	0.291737	Probability	0.589110	
ARCH(2) Test:				
F-statistic	4.389019	Probability	0.013304	
Obs*R-squared	8.597398	Probability	0.013586	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000670	0.000111	6.051737	0.0000
RESID^2(-1)	0.026561	0.059732	0.444680	0.6569
RESID^2(-2)	0.174070	0.059764	2.912628	0.0039
ARCH(4) Test:				
F-statistic	2.806997	Probability	0.026118	
Obs*R-squared	10.97756	Probability	0.026817	
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000595	0.000124	4.786558	0.0000
RESID^2(-1)	0.006613	0.061079	0.108263	0.9139
RESID^2(-2)	0.165760	0.060850	2.724060	0.0069
RESID^2(-3)	0.092096	0.060863	1.513182	0.1314
RESID^2(-4)	0.030711	0.061138	0.502325	0.6159



b) Estimation du modèle ARCH(2,0)

Sur EViews, suivre : Quick/Estimate Equation... → les boîtes de dialogue suivantes apparaissent (elles complètent la procédure : à gauche, choisir la méthode ARCH ; et à droite, choisir le type du modèle et préciser l'ordre du retard) :

Autrement :

Sur EViews, faire : arch(2,0) EURUS C EURUS(-1) EURUS(-2)

Dependent Variable: EURUS Method: ML - ARCH (Marquardt) - Normal distribution Date: 01/05/14 Time: 17:17 Sample (adjusted): 1990M03 2013M03 Included observations: 277 after adjustments Convergence achieved after 17 iterations Variance backcast: ON GARCH = C(4) + C(5)*RESID(-1)^2 + C(6)*RESID(-2)^2				
	Coefficient	Std. Error	z-Statistic	Prob.
C	0.028306	0.014212	1.991780	0.0464
EURUS(-1)	1.314995	0.060422	21.76340	0.0000
EURUS(-2)	-0.338019	0.060333	-5.602554	0.0000
Variance Equation				
C	0.000696	7.34E-05	9.491408	0.0000
RESID(-1)^2	0.022554	0.050829	0.443732	0.6572
RESID(-2)^2	0.136189	0.066153	2.058679	0.0395
R-squared	0.965826	Mean dependent var	1.216163	
Adjusted R-squared	0.965196	S.D. dependent var	0.156689	
S.E. of regression	0.029232	Akaike info criterion	-4.233367	
Sum squared resid	0.231570	Schwarz criterion	-4.154868	
Log likelihood	592.3213	F-statistic	1531.806	
Durbin-Watson stat	1.938197	Prob(F-statistic)	0.000000	

Date: 01/05/14 Time: 17:19 Sample: 1990M03 2013M03 Included observations: 277						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.035	0.035	0.3460	0.556
		2	-0.118	-0.119	4.2274	0.121
		3	0.077	0.087	5.9014	0.117
		4	0.001	-0.020	5.9019	0.207
		5	-0.063	-0.044	7.0367	0.218
		6	0.018	0.015	7.1343	0.309
		7	-0.087	-0.103	9.3062	0.231
		8	-0.006	0.017	9.3173	0.316
		9	0.042	0.015	9.8193	0.365
		10	0.001	0.011	9.8196	0.456
		11	-0.050	-0.044	10.534	0.483
		12	0.026	0.017	10.736	0.552
		13	0.009	-0.001	10.760	0.631
		14	-0.083	-0.080	12.775	0.544
		15	0.026	0.035	12.978	0.604



_____ Test d'effets ARCH sur le modèle ARCH(2,0) estimé : pas d'effets ARCH (prob>5%), le modèle ARCH(2,0) est accepté.

ARCH Test:			
F-statistic	0.047581	Probability	0.827490
Obs*R-squared	0.047920	Probability	0.826724

c) **Estimation du modèle AR(2) avec erreurs GARCH(2,0) : c'est le modèle qui nous a paru significatif.**

_____ Sur **EViews**, fixer « ARCH : 0 et GARCH(2) » : à droite, le corrélogramme des résidus simple.

Dependent Variable: EURUS Method: ML - ARCH (Marquardt) - Normal distribution Date: 01/05/14 Time: 17:24 Sample (adjusted): 1990M03 2013M03 Included observations: 277 after adjustments Convergence achieved after 67 iterations Variance backcast: ON GARCH = C(4) + C(5)*GARCH(-1) + C(6)*GARCH(-2)				
	Coefficient	Std. Error	z-Statistic	Prob.
C	0.030462	0.015262	1.995852	0.0459
EURUS(-1)	1.296141	0.057142	22.68270	0.0000
EURUS(-2)	-0.320899	0.055926	-5.737889	0.0000
Variance Equation				
C	4.83E-05	9.27E-06	5.208103	0.0000
GARCH(-1)	1.929195	0.032049	60.19537	0.0000
GARCH(-2)	-0.986483	0.027702	-35.61050	0.0000
R-squared	0.965838	Mean dependent var	1.216163	
Adjusted R-squared	0.965208	S.D. dependent var	0.156689	
S.E. of regression	0.029227	Akaike info criterion	-4.218793	
Sum squared resid	0.231487	Schwarz criterion	-4.140294	
Log likelihood	590.3028	F-statistic	1532.378	
Durbin-Watson stat	1.905273	Prob(F-statistic)	0.000000	

Date: 01/05/14 Time: 17:26 Sample: 1990M03 2013M03 Included observations: 277					
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.048	0.048	0.6367	0.425
		2 -0.137	-0.139	5.8862	0.053
		3 0.082	0.098	7.7886	0.051
		4 -0.007	-0.039	7.8042	0.099
		5 -0.053	-0.025	8.5986	0.126
		6 0.022	0.014	8.7406	0.189
		7 -0.097	-0.111	11.424	0.121
		8 -0.001	0.027	11.424	0.179
		9 0.063	0.027	12.557	0.184
		10 0.010	0.024	12.584	0.248
		11 -0.049	-0.043	13.270	0.276
		12 0.015	0.008	13.333	0.345
		13 0.014	0.004	13.390	0.418
		14 -0.081	-0.082	15.297	0.358
		15 0.037	0.055	15.712	0.401

Corrélogramme des résidus aux carrés

Date: 01/05/14 Time: 17:29 Sample: 1990M03 2013M03 Included observations: 277					
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.029	0.029	0.2369	0.626
		2 0.185	0.184	9.8746	0.007
		3 0.096	0.089	12.460	0.006
		4 0.070	0.034	13.838	0.008
		5 0.233	0.207	29.253	0.000
		6 -0.059	-0.094	30.238	0.000
		7 0.062	-0.021	31.336	0.000
		8 0.024	0.015	31.499	0.000
		9 0.102	0.090	34.521	0.000
		10 -0.023	-0.081	34.675	0.000
		11 0.019	0.023	34.784	0.000
		12 -0.026	-0.042	34.976	0.000
		13 -0.043	-0.061	35.508	0.001
		14 0.049	0.030	36.212	0.001
		15 -0.034	0.025	36.548	0.001



d) Estimation du modèle AR(2) avec erreurs TGARCH(2,1) : c'est le modèle qui nous a paru significatif.

Sur **EViews**, fixer « ARCH : 1, GARCH : 2, Threshold order : 1 » :

Dependent Variable: EURUS Method: ML - ARCH (Marquardt) - Normal distribution Date: 01/05/14 Time: 17:34 Sample (adjusted): 1990M03 2013M03 Included observations: 277 after adjustments Convergence achieved after 33 iterations Variance backcast: ON GARCH = C(4) + C(5)*RESID(-1)^2 + C(6)*RESID(-1)^2*(RESID(-1)<0) + C(7)*GARCH(-1) + C(8)*GARCH(-2)					Date: 01/05/14 Time: 17:36 Sample: 1990M03 2013M03 Included observations: 277					
	Coefficient	Std. Error	z-Statistic	Prob.	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
C	0.030524	0.012777	2.388933	0.0169			1 -0.042 -0.042	0.4918	0.483	
EURUS(-1)	1.287174	0.059580	21.60413	0.0000			2 0.043 0.041	1.0124	0.603	
EURUS(-2)	-0.311911	0.059679	-5.226487	0.0000			3 0.018 0.021	1.1028	0.776	
Variance Equation							4 -0.031 -0.031	1.3758	0.848	
C	6.86E-06	1.60E-06	4.277212	0.0000			5 0.133 0.129	6.4046	0.269	
RESID(-1)^2	0.013850	0.005656	2.448956	0.0143			6 -0.091 -0.080	8.7639	0.187	
RESID(-1)^2*(RESID(-1)<0)	-0.016190	0.006330	-2.557801	0.0105			7 0.013 -0.002	8.8131	0.266	
GARCH(-1)	1.932005	0.034659	55.74301	0.0000			8 0.003 0.006	8.8161	0.358	
GARCH(-2)	-0.945963	0.033070	-28.60447	0.0000			9 0.035 0.046	9.1757	0.421	
R-squared	0.965831	Mean dependent var	1.216163				10 -0.076 -0.099	10.854	0.369	
Adjusted R-squared	0.964942	S.D. dependent var	0.156689				11 -0.004 0.012	10.859	0.455	
S.E. of regression	0.029338	Akaike info criterion	-4.262507				12 -0.057 -0.062	11.808	0.461	
Sum squared resid	0.231536	Schwarz criterion	-4.157842				13 -0.102 -0.104	14.836	0.318	
Log likelihood	598.3572	F-statistic	1086.238				14 0.009 -0.008	14.858	0.388	
Durbin-Watson stat	1.889449	Prob(F-statistic)	0.000000				15 -0.044 -0.004	15.425	0.421	

Constat : Les résidus de l'estimation de ce modèle sont de bruits blancs : le processus AR(2) avec erreurs TGARCH(2,1) est accepté.

e) Estimation du modèle AR(2) avec erreurs EGARCH(2,1) : c'est le modèle qui nous a paru significatif.

Sur **EViews**, fixer « ARCH : 1, GARCH : 2, Asymetric order : 1 » (NB : dans "Model", choisir "EGARCH") :



Dependent Variable: EURUS Method: ML - ARCH (Marquardt) - Normal distribution Date: 01/05/14 Time: 17:43 Sample (adjusted): 1990M03 2013M03 Included observations: 277 after adjustments Convergence achieved after 56 iterations Variance backcast: ON LOG(GARCH) = C(4) + C(5)*ABS(RESID(-1)/@SQRT(GARCH(-1))) + C(6)*RESID(-1)/@SQRT(GARCH(-1)) + C(7)*LOG(GARCH(-1)) + C(8)*LOG(GARCH(-2))					Date: 01/05/14 Time: 17:46 Sample: 1990M03 2013M03 Included observations: 277						
					Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
							1	0.117	0.117	3.8404	0.050
							2	0.129	0.117	8.5187	0.014
							3	0.047	0.020	9.1291	0.028
							4	0.097	0.077	11.769	0.019
							5	0.211	0.192	24.382	0.000
							6	-0.048	-0.114	25.047	0.000
							7	0.055	0.024	25.927	0.001
							8	0.023	0.022	26.081	0.001
							9	0.116	0.080	29.983	0.000
							10	0.002	-0.057	29.984	0.001
							11	0.014	0.029	30.045	0.002
							12	-0.020	-0.048	30.167	0.003
							13	-0.045	-0.060	30.750	0.004
							14	0.026	0.011	30.942	0.006
							15	-0.026	0.008	31.146	0.008

Constat : Les résidus de l'estimation de ce modèle ne semblent pas de bruits blancs : le processus AR(2) avec erreurs EGARCH(2,1) n'est pas accepté. A cela s'ajoute que le coefficient $c(5)$ est négatif, mais statistiquement non significatif.

f) **Estimation du modèle AR(2) avec erreurs ARCH(2)-M :** c'est le modèle qui nous a paru significatif.

EViews offre deux possibilités pour estimer le modèle ARCH-M, à savoir :

- o 1^{ère} possibilité/Hypothèse : la variance conditionnelle intervient dans l'espérance conditionnelle comme suit (notre cas : AR(2) avec erreurs ARCH(2)-M) :

$$EURUS_t = a_0 + a_1 EURUS_{t-1} + a_2 EURUS_{t-2} + \lambda h_t^2 + e_t$$

$$h_t^2 = a_0 + a_1 e_{t-1}^2 + a_2 e_{t-2}^2$$

Sur **EViews**, suivre : Quick/Estimate Equation... → la boîte de dialogue à gauche apparaît (elle complète la procédure) : Estimer le processus AR(2) sans constante avec erreurs ARCH(2)-M ci-dessus : 1^{ère} possibilité – variance.



Dependent Variable: EURUS
Method: ML - ARCH (Marquardt) - Normal distribution
Date: 01/12/14 Time: 17:40
Sample (adjusted): 1990M03 2013M03
Included observations: 277 after adjustments
Convergence achieved after 16 iterations
Variance backcast: OFF
GARCH = C(4) + C(5)*RESID(-1)^2 + C(6)*RESID(-2)^2

	Coefficient	Std. Error	z-Statistic	Prob.
GARCH	37.21625	6.782951	5.486735	0.0000
EURUS(-1)	1.308725	0.033280	39.32504	0.0000
EURUS(-2)	-0.332028	0.031868	-10.41883	0.0000

Variance Equation

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000802	6.52E-05	12.29029	0.0000
RESID(-1)^2	-0.073855	0.028884	-2.556920	0.0106
RESID(-2)^2	0.076318	0.028730	2.656362	0.0079

R-squared	0.966493	Mean dependent var	1.216163
Adjusted R-squared	0.965874	S.D. dependent var	0.156689
S.E. of regression	0.028945	Akaike info criterion	-4.265261
Sum squared resid	0.227054	Schwarz criterion	-4.186763
Log likelihood	596.7387	Durbin-Watson stat	1.907045

- o 2^{ème} possibilité/Hypothèse : l'écart-type conditionnel intervient dans l'espérance conditionnelle comme suit (notre cas : AR(2) avec erreurs ARCH(2)-M)

$$EURUS_t = a_0 + a_1 EURUS_{t-1} + a_2 EURUS_{t-2} + \lambda h_t + e_t$$

$$h_t^2 = a_0 + a_1 e_{t-1}^2 + a_2 e_{t-2}^2$$

Sur **EViews** (même procédure que le modèle précédent) : Estimer le processus AR(2) sans constante avec erreurs ARCH(2)-M ci-dessus : 2^{ème} possibilité – écart-type (dans « ARCH-M » cocher : Std. Dev.) :



Dependent Variable: EURUS				
Method: ML - ARCH (Marquardt) - Normal distribution				
Date: 01/12/14 Time: 17:44				
Sample (adjusted): 1990M03 2013M03				
Included observations: 277 after adjustments				
Convergence achieved after 10 iterations				
Variance backcast: OFF				
GARCH = C(4) + C(5)*RESID(-1)^2 + C(6)*RESID(-2)^2				
	Coefficient	Std. Error	z-Statistic	Prob.
@SQRT(GARCH)	1.464858	0.277149	5.285453	0.0000
EURUS(-1)	1.285147	0.014670	87.60177	0.0000
EURUS(-2)	-0.317559	0.017498	-18.14884	0.0000
Variance Equation				
C	0.000789	6.85E-05	11.51605	0.0000
RESID(-1)^2	-0.079056	0.020678	-3.823227	0.0001
RESID(-2)^2	0.120508	0.051430	2.343137	0.0191
R-squared	0.966248	Mean dependent var	1.216163	
Adjusted R-squared	0.965626	S.D. dependent var	0.156689	
S.E. of regression	0.029051	Akaike info criterion	-4.254145	
Sum squared resid	0.228709	Schwarz criterion	-4.175646	
Log likelihood	595.1991	Durbin-Watson stat	1.860842	

g) Estimation du modèle AR(2) avec erreurs GARCH(2,0)-M : c'est le modèle qui nous a paru significatif.

EViews offre deux possibilités pour estimer le modèle GARCH-M, à savoir :

- o 1^{ère} possibilité/Hypothèse : la variance conditionnelle intervient dans l'espérance conditionnelle comme suit (notre cas : AR(2) avec erreurs GARCH(2)-M)

$$EURUS_t = a_0 + a_1 EURUS_{t-1} + a_2 EURUS_{t-2} + \lambda h_t^2 + e_t$$

$$h_t^2 = a_0 + a_1 e_{t-1}^2 + a_2 e_{t-2}^2 + b_1 h_{t-1}^2 + b_2 h_{t-2}^2$$

____ Sur **EViews**, suivre : Quick/Estimate Equation... → la boîte de dialogue à gauche apparaît (elle complète la procédure) : Estimer le processus « AR(2) sans constante avec erreurs GARCH(2)-M » ci-dessus : 1^{ère} possibilité – variance.



Dependent Variable: EURUS
Method: ML - ARCH (Marquardt) - Normal distribution
Date: 01/12/14 Time: 17:49
Sample (adjusted): 1990M03 2013M03
Included observations: 277 after adjustments
Convergence not achieved after 500 iterations
Variance backcast: OFF
GARCH = C(4) + C(5)*GARCH(-1) + C(6)*GARCH(-2)

	Coefficient	Std. Error	z-Statistic	Prob.
GARCH	39.21532	16.49258	2.377755	0.0174
EURUS(-1)	1.300398	0.051522	25.23975	0.0000
EURUS(-2)	-0.325476	0.050664	-6.424162	0.0000

Variance Equation				
C	0.000609	4.10E-05	14.85724	0.0000
GARCH(-1)	1.218890	0.008461	144.0583	0.0000
GARCH(-2)	-0.999246	0.010398	-96.09606	0.0000

R-squared	0.965979	Mean dependent var	1.216163
Adjusted R-squared	0.965351	S.D. dependent var	0.156689
S.E. of regression	0.029167	Akaike info criterion	-4.215245
Sum squared resid	0.230536	Schwarz criterion	-4.136746
Log likelihood	589.8114	Durbin-Watson stat	1.917752

- o 2^{ème} possibilité/Hypothèse : l'écart-type conditionnel intervient dans l'espérance conditionnelle comme suit (notre cas : AR(2) avec erreurs GARCH(2)-M) :

$$EURUS_t = a_0 + a_1 EURUS_{t-1} + a_2 EURUS_{t-2} + \lambda h_t + e_t$$

$$h_t^2 = a_0 + a_1 e_{t-1}^2 + a_2 e_{t-2}^2 + b_1 h_{t-1}^2 + b_2 h_{t-2}^2$$

Sur **EViews** (même procédure que le modèle précédent) : Estimer le processus AR(2) sans constante avec erreurs GARCH(2)-M : 2^{ème} possibilité – Ecart-type (dans « ARCH-M » cocher : Std. Dev.) :

Dependent Variable: EURUS				
Method: ML - ARCH (Marquardt) - Normal distribution				
Date: 01/12/14 Time: 17:53				
Sample (adjusted): 1990M03 2013M03				
Included observations: 277 after adjustments				
Convergence not achieved after 500 iterations				
Variance backcast: OFF				
GARCH = C(4) + C(5)*GARCH(-1) + C(6)*GARCH(-2)				
	Coefficient	Std. Error	z-Statistic	Prob.
@SQRT(GARCH)	1.142573	0.478340	2.388621	0.0169
EURUS(-1)	1.296533	0.050654	25.59565	0.0000
EURUS(-2)	-0.322840	0.049908	-6.468636	0.0000
Variance Equation				
C	0.000553	4.18E-05	13.25036	0.0000
GARCH(-1)	1.302960	0.007699	169.2473	0.0000
GARCH(-2)	-1.003735	0.010286	-97.57954	0.0000
R-squared	0.965970	Mean dependent var	1.216163	
Adjusted R-squared	0.965342	S.D. dependent var	0.156689	
S.E. of regression	0.029170	Akaike info criterion	-4.213923	
Sum squared resid	0.230595	Schwarz criterion	-4.135425	
Log likelihood	589.6284	Durbin-Watson stat	1.909731	



2) Recherche du modèle optimal

Le tableau ci-dessous présente les résultats des différents modèles estimés (nous y insérons également quelques critères de sélection d'un modèle optimal) :

	ARCH(2)	GARCH(2,0)	TGARCH(2,1)	EGARCH(2,1)	ARCH(2)-M Variance	ARCH(2)-M Ecart-type	GARCH(2,0)-M Variance	GARCH(2,0)-M Ecart-type
Equation de la Moyenne								
Constante (Prob)	0.03 (0.046)	0.03 (0.046)	0.07 (0.02)	0.04 (0.02)	*	*	*	*
AR(1)	1.31 (0.000)	1.30 (0.00)	1.29 (0.00)	1.36 (0.00)	1.31 (0.000)	1.29 (0.00)	1.3 (0.00)	1.30 (0.00)
AR(2)	-0.33 (0.000)	-0.32 (0.00)	-0.31 (0.00)	-0.39 (0.00)	-0.33 (0.000)	-0.32 (0.00)	-0.33 (0.00)	-0.32 (0.00)
Equation de la Variance								
Constante	0.0007 (0.000)	0.005 (0.00)	0.0007 (0.00)	-18.37 (0.00)	0.0008 (0.000)	0.0008 (0.00)	0.0006 (0.00)	0.0006 (0.00)
ARCH(1)	0.02 (0.68)	*	0.01 (0.01)	*	-0.07 (0.01)	-0.08 (0.00)	*	*
ARCH(2)	0.14 (0.04)	*	*	*	0.08 (0.008)	-0.12 (0.02)	*	*
GARCH(1)	*	1.93 (0.00)	1.93 (0.00)	-0.79 (0.00)	*	*	1.22 (0.00)	1.30 (0.00)
GARCH(2)	*	-0.99 (0.00)	-0.95 (0.00)	-0.81 (0.00)	*	*	-0.999 (0.00)	-1.004 (0.00)
RESID(-1)^2*(RESID(-1)<0)	*	*	-0.02 (0.01)	*	*	*	*	*



	ARCH(2)	GARCH(2,0)	TGARCH(2,1)	EGARCH(2,1)	ARCH(2)-M Variance	ARCH(2)-M Ecart-type	GARCH(2,0)-M Variance	GARCH(2,0)-M Ecart-type
ABS(RESID(-1)/ @SQRT/GARCH(-1)))	*	*	*	-0.19 (0.07)	*	*	*	*
(RESID(-1)/ @SQRT/GARCH(-1))	*	*	*	0.09 (0.17)	*	*	*	*
GARCH	*	*	*	*	37.22 (0.000)	*	39.22 (0.02)	*
@SQRT(GARCH)	*	*	*	*	*	1.46 (0.00)	*	1.14 (0.02)
R²	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
AIC	-4.23	-4.22	-4.26	-4.22	-4.27	-4.25	-4.22	-4.21
DW	1.94	1.91	1.89	2.003	1.91	1.86	1.92	1.91
Ljung-Box/LB(36) (Prob)	27.83 (0.83)	32.35 (0.64)	27.17 (0.86)	28.23 (0.82)	28.92 (0.79)	27.32 (0.85)	34.13 (0.56)	33.62 (0.58)
LB²(36) (Prob)	40.168 (0.29)	56.01 (0.02)	22.67 (0.96)	44.38 (0.16)	50.94 (0.05)	46.59 (0.11)	57.02 (0.01)	60.37 (0.007)
Normalité	Non	Non	Oui	Non	Oui	Oui	Non	Non
Effet ARCH	Non	Oui(lag :2)	Non	Oui(lag :2)	Non	Non	Oui(lag :2)	Oui(lag :2)

_____ **Commentaire** : Si l'on s'en tient au corrélogramme des résidus aux carrés (Cfr **LB²(36)**), il apparaît que – pour quatre modèles/processus – les résidus issus de la variance conditionnelle ne sont pas des bruits blancs (prob<5%). Il s'agit des modèles : GARCH(2,0) ; ARCH(2)-M/Variance ; GARCH(2,0)-M/Variance et GARCH(2,0)-M/Ecart-type.

_____ **Modèle optimal** : au regard des résultats ci-dessus, l'on note le modèle optimal – selon les critères – comme suit :

- o AIC minimum (qualité de l'équation de la moyenne) : **ARCH(2)-M/Variance** ;
- o DW maximal : **EGARCH(2,1)** ;
- o LB²(36) minimum (qualité de l'équation de la moyenne) : **TGARCH(2,1)** ;



Ainsi, trois modèles sont candidats à la prévision. Dans le point qui suit, nous nous servons des critères de bonne prévision (Mean Absolute Percentage Error/MAPE minimal et Coefficient d'inégalité de Theil proche de zéro) pour sélectionner le modèle optimal parmi les trois.

E. Prévision

a) Choix du modèle optimal (recours aux critères de bonne prédiction)

Nous recourons aux procédures de simulation dynamique de « EViews » pour prévoir notre série brute « EURUS » à travers nos huit modèles (ne disposant pas de nouvelles données, notre série brute a été tronquée de 5 mois. Sur base de cette série tronquée, une ré-estimation nous a aidée à obtenir les prédictions souhaitées).

_____ Pour ce faire, sur EViews, procéder comme suit :

- o Taper : `smp1 1990:01 2012:10`
- o Estimer les paramètres du modèle sélectionné ;
- o après l'estimation des paramètres, dans l'output des résultats, cliquer sur **Forecast** : la boîte de dialogue suivante apparaît (elle complète la procédure.
Ex : ARCH(2)-M avec Variance) → cliquer sur « Ok » :

_____ Les valeurs MAPE et Theil recueillies, pour différents modèles (nous intégrons aussi les valeurs AIC/Akaike Criteria), sont favorables pour le processus « **ARCH(2)-M/Variance** » estimé (voir le tableau ci-dessous). Ce dernier est ainsi retenu pour la prévision des valeurs futures de la parité « Euro-Usd » (de Novembre 2012 à Mars 2013, soit 5 mois).

Critères	Modèles/Processus			
	ARCH(2)	GARCH(2,0)	TGARCH(2,1)	EGARCH(2,1)
MAPE	1.4207	1.4463	1.4363	1.4960
THEIL	0.0079	0.0083	0.0081	0.0091
AIC	*	*	*	*



Critères	Modèles/Processus (suite)			
	ARCH(2)-M/ Variance	ARCH(2)-M/ Ecart-type	GARCH(2,0)-M/ Variance	GARCH(2,0)-M/ Ecart-type
MAPE	1.4125	1.4151	1.6857	1.9980
THEIL	0.0079	0.0078	0.010	0.0116
AIC	-4.2616	-4.2575	*	*

_____ **Modèle optimal** : selon les critères, l'on note le modèle optimal comme suit :

- o MAPE : **ARCH(2)-M/Variance** ;
- o THEIL : **ARCH(2)-M/Ecart-type** ;
- o AIC minimum (qualité de l'équation de la moyenne) : **ARCH(2)-M/Variance**.

Ce qui revient à dire que le modèle « **ARCH(2)-M/Variance** » est optimal/meilleur (il remplit les propriétés statistiques d'une bonne prévision).

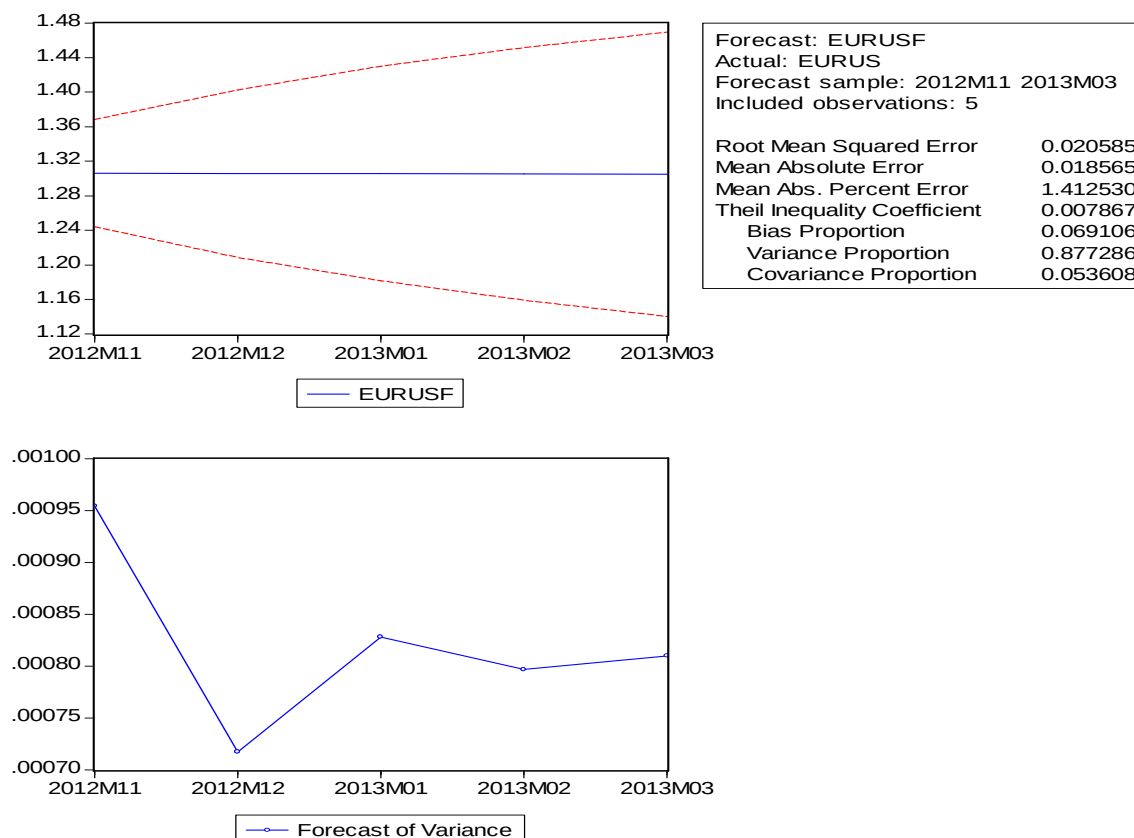
_____ **NB** : les résultats du modèle « ARCH(2)-M avec Variance » estimé sur la période « 1990:01 à 2012:10 » se présentent comme suit (Voir « AIC ») :

Dependent Variable: EURUS				
Method: ML - ARCH (Marquardt) - Normal distribution				
Date: 01/17/14 Time: 19:37				
Sample (adjusted): 1990M03 2012M10				
Included observations: 272 after adjustments				
Convergence achieved after 18 iterations				
Variance backcast: OFF				
GARCH = C(4) + C(5)*RESID(-1)^2 + C(6)*RESID(-2)^2				
	Coefficient	Std. Error	z-Statistic	Prob.
GARCH	34.76198	6.694769	5.192408	0.0000
EURUS(-1)	1.299030	0.036025	36.05941	0.0000
EURUS(-2)	-0.320755	0.034476	-9.303691	0.0000
Variance Equation				
C	0.000786	6.48E-05	12.13614	0.0000
RESID(-1)^2	-0.072319	0.035717	-2.024805	0.0429
RESID(-2)^2	0.098253	0.037264	2.636641	0.0084
R-squared	0.966628	Mean dependent var	1.214419	
Adjusted R-squared	0.966001	S.D. dependent var	0.157570	
S.E. of regression	0.029054	Akaike info criterion	-4.261566	
Sum squared resid	0.224542	Schwarz criterion	-4.182026	
Log likelihood	585.5729	Durbin-Watson stat	1.879145	

b) Prévision

_____ Cfr procédure EViews (Choix du modèle optimal) : Cliquer sur « Ok » pour obtenir les graphiques suivants – ils concernent « ARCH(2)-M avec Variance » – d'où nous avons tiré les valeurs « MAPE » et « Theil » :

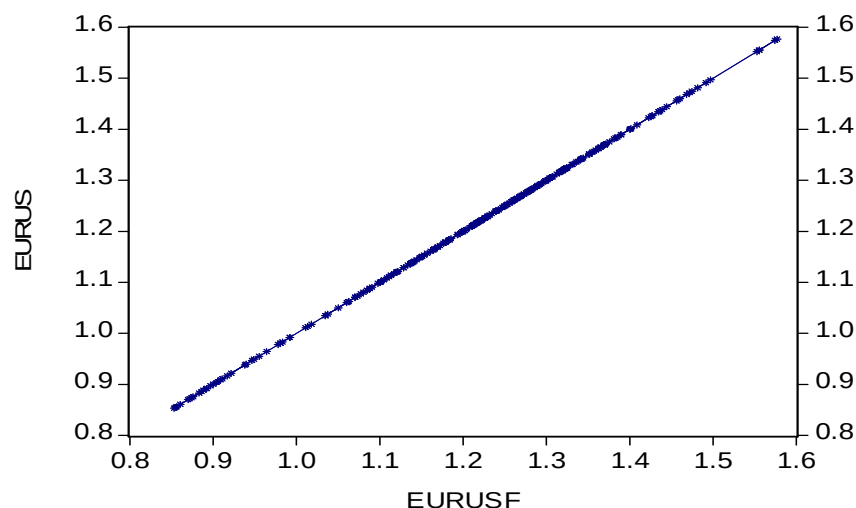


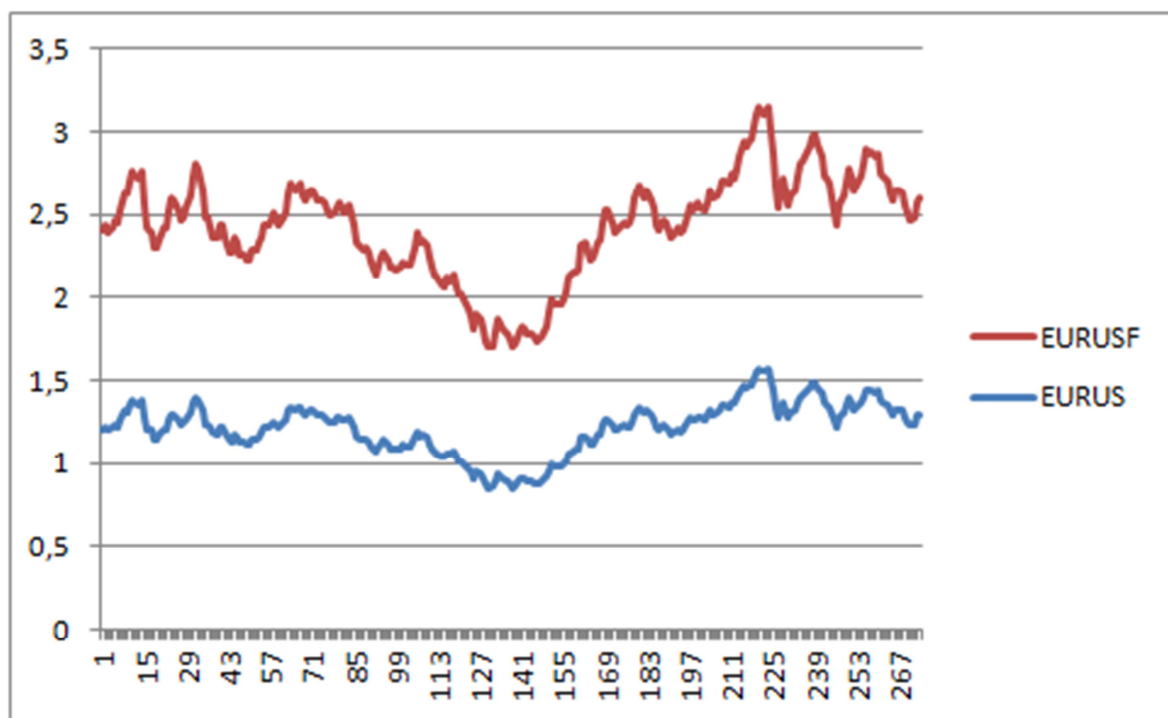
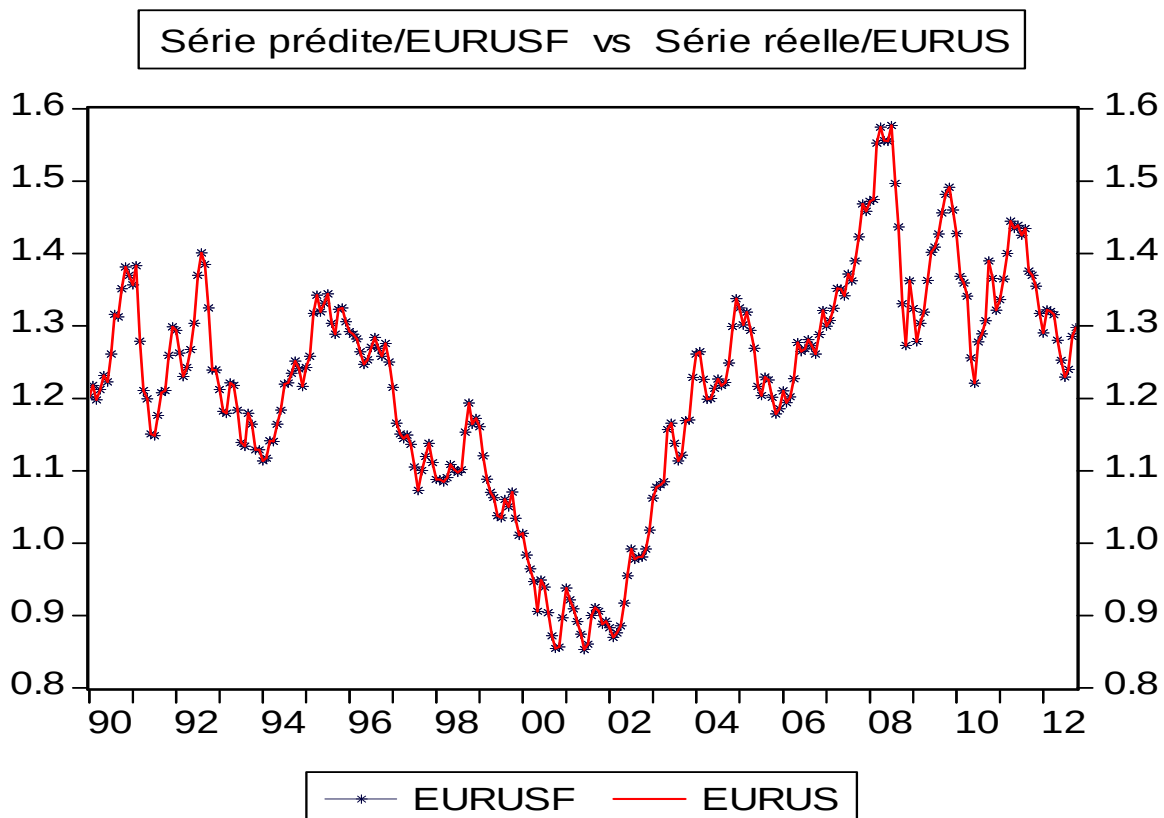


Obtenir les données réelles/EURUS et prédites/EURUSF en cliquant sur chaque observation (Cfr fichier de travail) :

Mois	EURUS	EURUSF
2012M11	1,282694124	1,306243916
2012M12	1,312795379	1,305643272
2013M01	1,328228317	1,305873768
2013M02	1,335710899	1,305277312
2013M03	1,295552553	1,304885629

Produire les graphiques traduisant l'évolution des données réelles et prédites (sur EViews et Excel) : les graphiques montrent que la prédiction est presque parfaite.





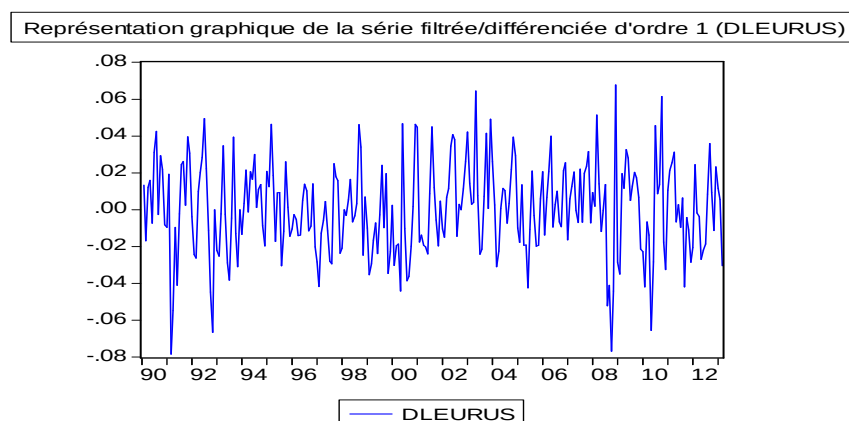
VII.2.3. ETUDE DE LA SÉRIE FILTRÉE « DLEURUS » PAR LA MODELISATION HETEROSCEDASTIQUE (ARCH)

_____ **NB** : sur **EViews**, la série filtrée (différenciée d'ordre 1 : Cfr procédure Box et Jenkins) est obtenue en faisant : $genr\ DLEURUS=d(log(EURUS))$

_____ Nous allons estimer quatre modèles (ARCH, GARCH, TGARCH et EGARCH) parmi lesquels nous retiendrons le plus performant, et allons procéder comme suit :

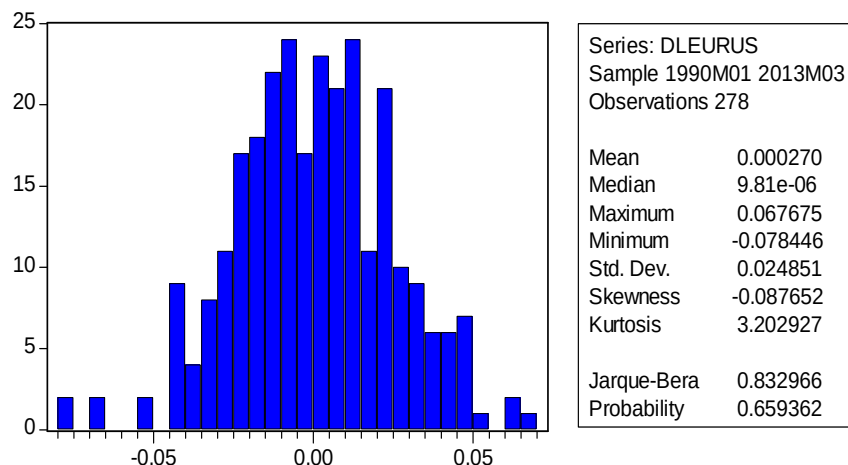
- (i) Représentation graphique de la série filtrée « DLEURUS » ;
- (ii) Statistiques descriptives et test de normalité de la série « DLEURUS » ;
- (iii) Etude de la volatilité de la série « DLEURUS » ;
- (iv) Test de bruit blanc de la série filtrée « DLEURUS » ;
- (v) Estimation des modèles (4) et recherche du modèle optimal
- (vi) Prévision

A. Représentation graphique de la série filtrée « DLEURUS »



_____ **Constat** : la série filtrée « DLEURUS » est stationnaire en moyenne, mais non stationnaire en variance (elle est très volatile : l'on peut même distinguer des groupements des fortes variations ou des faibles variations sur la série) : ce qui justifie le recours à la modélisation hétéroscédastique pour étudier notre série filtrée. **NB** : l'échantillon a 278 observations (279 – 1).

B. Statistiques descriptives et test de normalité de la série filtrée « DLEURUS »



_____ **Constats :** Au regard de la statistique de Jarque-Berra ($\text{prob} > 5\%$), notre série filtrée est normalement distribuée. Toutefois, le coefficient de Kurtosis ($3.203 > 3$) et celui de Skewness ($-0.088 < 0$) calculés nous renseignent respectivement sur la possibilité d'occurrence d'observations extrêmes et sur le fait que notre paire « EURUS » stationnarisée réagit plus à un choc négatif que positif (la distribution est étalée vers la gauche).

C. Etude de la volatilité de la série filtrée « DLEURUS »

Comme pour la série brute, régressons « DLEURUS » au carré sur cette même série décalée (cette façon de procéder permet de juger de l'autocorrélation de la variance des résidus/hétéroscédasticité).

Dependent Variable: DLEURUS2				
Method: Least Squares				
Date: 01/16/14 Time: 12:25				
Sample (adjusted): 1990M03 2013M03				
Included observations: 277 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000531	6.57E-05	8.076919	0.0000
DLEURUS2(-1)	0.139999	0.059696	2.345184	0.0197
R-squared	0.019607	Mean dependent var		0.000617
Adjusted R-squared	0.016042	S.D. dependent var		0.000916
S.E. of regression	0.000908	Akaike info criterion		-11.16310
Sum squared resid	0.000227	Schwarz criterion		-11.13693
Log likelihood	1548.090	F-statistic		5.499888
Durbin-Watson stat	2.027573	Prob(F-statistic)		0.019729

_____ **Constat :** le coefficient associé à « DLEURUS2(-1) » est statistiquement significatif, ce qui permet d'accepter l'hypothèse d'hétéroscédasticité conditionnelle (la variation de « EURUS » au temps « t » est fonction de son évolution au temps « t-1 »).

D. Test de bruit blanc de la série filtrée « DLEURUS »

Nous nous servons du corrélogramme ci-dessous pour signaler que notre série filtrée présente au moins 1 terme significatif de part et d'autre (Cfr méthodologie Box et Jenkins, « DLEURUS » est identifiée comme un *arma* (1,0,1), un *ar*(1), soit un *ma*(1)) et, de ce fait, elle n'est pas un bruit blanc ($\text{prob} < 5\%$).



Date: 01/16/14 Time: 12:04 Sample: 1990M01 2013M03 Included observations: 278						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.312	0.312	27.335	0.000
		2	-0.033	-0.144	27.634	0.000
		3	0.021	0.088	27.764	0.000
		4	-0.042	-0.094	28.255	0.000
		5	-0.065	-0.014	29.462	0.000
		6	-0.031	-0.018	29.736	0.000
		7	-0.081	-0.080	31.621	0.000
		8	-0.011	0.050	31.658	0.000
		9	0.042	0.010	32.171	0.000
		10	0.007	-0.005	32.186	0.000
		11	-0.046	-0.056	32.793	0.001
		12	-0.027	-0.003	33.009	0.001
		13	-0.046	-0.048	33.629	0.001
		14	-0.090	-0.070	36.025	0.001
		15	0.011	0.067	36.060	0.002

E. Estimation des modèles et recherche du modèle optimal

1) Estimation des modèles

(i) Estimation du modèle ARCH

Nous estimons un $AR(1)^{(16)}$ ou un $AR(2)$ à partir duquel nous effectuons le test de présence d'effets ARCH pour juger de la nécessité d'une modélisation ARCH.

► Estimation du modèle $AR(1)$:

Dependent Variable: DLEURUS Method: Least Squares Date: 01/16/14 Time: 13:32 Sample (adjusted): 1990M03 2013M03 Included observations: 277 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000104	0.001423	0.072813	0.9420
DLEURUS(-1)	0.312671	0.057424	5.444921	0.0000
R-squared	0.097316	Mean dependent var		0.000223
Adjusted R-squared	0.094034	S.D. dependent var		0.024883
S.E. of regression	0.023685	Akaike info criterion		-4.640788
Sum squared resid	0.154264	Schwarz criterion		-4.614622
Log likelihood	644.7492	F-statistic		29.64717
Durbin-Watson stat	1.895575	Prob(F-statistic)		0.000000

¹⁶ Le processus générateur de la série « EURUS » était un $arima(1,1,1)$ si l'on s'en tient aux résultats issus de la procédure Box-Jenkins appliquée à cette série. C'est ainsi que nous estimons le modèle $AR(1)$, le processus $arima(1,1,1)$ étant non significatif.



► Estimation du modèle AR(2) :

Dependent Variable: DLEURUS				
Method: Least Squares				
Date: 01/17/14 Time: 00:19				
Sample (adjusted): 1990M04 2013M03				
Included observations: 276 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLEURUS(-1)	0.360574	0.059874	6.022222	0.0000
DLEURUS(-2)	-0.147063	0.059848	-2.457270	0.0146
R-squared	0.117787	Mean dependent var	0.000285	
Adjusted R-squared	0.114567	S.D. dependent var	0.024907	
S.E. of regression	0.023437	Akaike info criterion	-4.661783	
Sum squared resid	0.150506	Schwarz criterion	-4.635548	
Log likelihood	645.3261	Durbin-Watson stat	1.960152	

_____ **Commentaire :** Nous retenons le processus AR(2) sans dérive qui modélise mieux la série filtrée que le processus AR(1) – l'on peut se servir aussi des critères Akaike et Schwarz (le processus AR(2) les minimise).

► Inférences sur le modèle AR(2) estimé :

_____ La statistique Q de Ljung-Box indique un terme statistiquement différent de zéro au 5^{ème} décalage (Cfr corrélogramme des résidus aux carrés) ; *Ce qui amène à présumer la présence d'effets ARCH d'ordre 5* dans notre série filtrée « DLEURUS » (le test ARCH ci-dessous est plus éloquent).

_____ *Test de présence d'effets ARCH sur le modèle AR(1) estimé : la probabilité critique étant < 5% pour un ARCH(5), nous acceptons l'hypothèse de présence d'effets ARCH d'ordre 5 (pour les lags 6,7..., les paramètres des termes autorégressifs se sont révélés non significatifs). D'où nous estimons que notre série filtrée suit un processus ARCH d'ordre 5.*

ARCH(5) Test :				
F-statistic	2.425079	Probability	0.035860	
Obs*R-squared	11.85738	Probability	0.036795	
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000373	8.49E-05	4.392651	0.0000
RESID^2(-1)	0.024061	0.060870	0.395287	0.6929
RESID^2(-2)	0.109087	0.060792	1.794437	0.0739
RESID^2(-3)	0.082113	0.060980	1.346564	0.1793
RESID^2(-4)	-0.036306	0.060799	-0.597150	0.5509
RESID^2(-5)	0.134511	0.060831	2.211217	0.0279

_____ **Note :** Signalons – pour les estimations qui suivent – que tous les processus estimés sont valides (les résidus issus des estimations sont tous normalement distribués), malgré la persistance d'effets ARCH pour les processus ARCH(5,0) et TGARCH(5,3).



(ii) **Estimation du modèle ARCH(5,0)** : ce processus est valide. NB : ARCH(5), GARCH(0) et Threshold order(1).

Dependent Variable: DLEURUS Method: ML - ARCH (Marquardt) - Normal distribution Date: 01/17/14 Time: 00:48 Sample (adjusted): 1990M04 2013M03 Included observations: 276 after adjustments Convergence achieved after 20 iterations Variance backcast: ON GARCH = C(3) + C(4)*RESID(-1)^2 + C(5)*RESID(-2)^2 + C(6)*RESID(-3)^2 + C(7)*RESID(-4)^2 + C(8)*RESID(-5)^2				
	Coefficient	Std. Error	z-Statistic	Prob.
DLEURUS(-1)	0.376872	0.062386	6.040972	0.0000
DLEURUS(-2)	-0.159946	0.060929	-2.625143	0.0087
Variance Equation				
C	0.000372	8.02E-05	4.637035	0.0000
RESID(-1)^2	0.021473	0.059046	0.363659	0.7161
RESID(-2)^2	0.123515	0.055599	2.221519	0.0263
RESID(-3)^2	0.029501	0.053975	0.546562	0.5847
RESID(-4)^2	-0.052399	0.060961	-0.859549	0.3900
RESID(-5)^2	0.203577	0.089737	2.268604	0.0233
R-squared	0.117488	Mean dependent var		0.000285
Adjusted R-squared	0.094438	S.D. dependent var		0.024907
S.E. of regression	0.023702	Akaike info criterion		-4.671300
Sum squared resid	0.150557	Schwarz criterion		-4.566360
Log likelihood	652.6393	Durbin-Watson stat		1.990322

(iii) **Estimation du modèle AR(2) sans dérive avec erreurs GARCH(4,0)** : c'est le modèle qui nous a paru significatif (il est valide). NB : ARCH(0) et GARCH(4).

Dependent Variable: DLEURUS Method: ML - ARCH (Marquardt) - Normal distribution Date: 01/17/14 Time: 01:02 Sample (adjusted): 1990M04 2013M03 Included observations: 276 after adjustments Convergence achieved after 36 iterations Variance backcast: ON GARCH = C(3) + C(4)*GARCH(-1) + C(5)*GARCH(-2) + C(6)*GARCH(-3) + C(7)*GARCH(-4)				
	Coefficient	Std. Error	z-Statistic	Prob.
DLEURUS(-1)	0.354977	0.056865	6.242460	0.0000
DLEURUS(-2)	-0.138599	0.051327	-2.700323	0.0069
Variance Equation				
C	0.000110	9.60E-05	1.147709	0.2511
GARCH(-1)	0.580696	0.229499	2.530280	0.0114
GARCH(-2)	0.495953	0.102373	4.844553	0.0000
GARCH(-3)	0.712997	0.084201	8.467803	0.0000
GARCH(-4)	-0.989256	0.235598	-4.198908	0.0000
R-squared	0.117714	Mean dependent var		0.000285
Adjusted R-squared	0.098035	S.D. dependent var		0.024907
S.E. of regression	0.023655	Akaike info criterion		-4.654539
Sum squared resid	0.150519	Schwarz criterion		-4.562717
Log likelihood	649.3264	Durbin-Watson stat		1.950572



- (iv) **Estimation du modèle AR(2) sans dérive avec erreurs TGARCH(5,3)** : c'est le modèle qui nous a paru significatif (il est valide). NB : ARCH(3), GARCH(5) et Threshold order(1).

Dependent Variable: DLEURUS Method: ML - ARCH (Marquardt) - Normal distribution Date: 01/17/14 Time: 01:14 Sample (adjusted): 1990M04 2013M03 Included observations: 276 after adjustments Convergence achieved after 52 iterations Variance backcast: ON GARCH = C(3) + C(4)*RESID(-1)^2 + C(5)*RESID(-2)^2 + C(6)*RESID(-3)^2 + C(7)*RESID(-1)^2*(RESID(-1)<0) + C(8)*GARCH(-1) + C(9)*GARCH(-2) + C(10)*GARCH(-3) + C(11)*GARCH(-4) + C(12)*GARCH(-5)				
	Coefficient	Std. Error	z-Statistic	Prob.
DLEURUS(-1)	0.368325	0.049798	7.396422	0.0000
DLEURUS(-2)	-0.194681	0.053040	-3.670435	0.0002
Variance Equation				
C	0.000106	5.84E-05	1.809176	0.0704
RESID(-1)^2	-0.102408	0.044068	-2.323868	0.0201
RESID(-1)^2*(RESID(-1)<0)	0.099577	0.029600	3.364050	0.0008
RESID(-2)^2	0.063756	0.029181	2.184880	0.0289
RESID(-3)^2	0.025779	0.035411	0.727995	0.4666
GARCH(-1)	1.009715	0.230055	4.389021	0.0000
GARCH(-2)	-0.121980	0.140223	-0.869903	0.3844
GARCH(-3)	0.534610	0.065001	8.224632	0.0000
GARCH(-4)	-1.152146	0.174603	-6.598654	0.0000
GARCH(-5)	0.504851	0.216804	2.328604	0.0199
R-squared	0.115697	Mean dependent var	0.000285	
Adjusted R-squared	0.078851	S.D. dependent var	0.024907	
S.E. of regression	0.023905	Akaike info criterion	-4.678550	
Sum squared resid	0.150863	Schwarz criterion	-4.521141	
Log likelihood	657.6399	Durbin-Watson stat	1.966548	

- (v) **Estimation du modèle AR(2) sans dérive avec erreurs EGARCH(4,1)** : c'est le modèle qui nous a paru significatif (il est valide). NB : ARCH(1), GARCH(4) et Asymmetric order(1).

Dependent Variable: DLEURUS Method: ML - ARCH (Marquardt) - Normal distribution Date: 01/17/14 Time: 01:36 Sample (adjusted): 1990M04 2013M03 Included observations: 276 after adjustments Convergence achieved after 112 iterations Variance backcast: ON LOG(GARCH) = C(3) + C(4)*ABS(RESID(-1)/@SQRT(GARCH(-1))) + C(5)*RESID(-1)/@SQRT(GARCH(-1)) + C(6)*LOG(GARCH(-1)) + C(7)*LOG(GARCH(-2)) + C(8)*LOG(GARCH(-3)) + C(9)*LOG(GARCH(-4))				
	Coefficient	Std. Error	z-Statistic	Prob.
DLEURUS(-1)	0.332892	0.042317	7.866658	0.0000
DLEURUS(-2)	-0.198577	0.039599	-5.014694	0.0000
Variance Equation				
C(3)	-59.30850	2.142623	-27.68033	0.0000
C(4)	-0.140267	0.053983	-2.598364	0.0094
C(5)	0.016009	0.021076	0.759557	0.4475
C(6)	-1.773905	0.064309	-27.58393	0.0000
C(7)	-2.461994	0.096577	-25.49247	0.0000
C(8)	-1.704552	0.087596	-19.45917	0.0000
C(9)	-0.899703	0.067469	-13.33496	0.0000
R-squared	0.113485	Mean dependent var	0.000285	
Adjusted R-squared	0.086923	S.D. dependent var	0.024907	
S.E. of regression	0.023800	Akaike info criterion	-4.688926	
Sum squared resid	0.151240	Schwarz criterion	-4.570870	
Log likelihood	656.0718	Durbin-Watson stat	1.891575	



2) Recherche du modèle optimal

Le tableau ci-dessous présente les résultats des différents modèles estimés (nous y insérons également quelques critères de sélection d'un modèle optimal) :

	ARCH(5)	GARCH(4,0)	TGARCH(5,3)	EGARCH(4,1)
Equation de la Moyenne				
Constante	*	*	*	*
AR(1) [z-stat] (Prob)	0.38 [6.04] (0.000)	0.35 [6.24] (0.00)	0.37 [7.40] (0.00)	0.33 [7.87] (0.00)
AR(2)	-0.16 [-2.63] (0.00)	-0.14 [-2.70] (0.007)	-0.19 [-3.67] (0.0002)	-0.20 [-5.01] (0.00)
Equation de la Variance				
Constante	0.0004 [4.64] (0.00)	0.0001 [1.15] (0.25)	*	-59.31 [-27.68] (0.00)
ARCH(1)	0.02 [0.36] (0.72)	*	-0.10 [-2.32] (0.02)	-1.77 [-27.58] (0.00)
ARCH(2)	0.12 [2.22] (0.03)	*	0.06 [2.18] (0.03)	-2.46 [-25.49] (0.00)
ARCH(3)	0.03 [0.55] (0.58)	*	0.03 [0.73] (0.47)	-1.70 [-19.46] (0.00)
ARCH(4)	-0.05 [-0.86] (0.39)	*	*	-0.90 [-13.33] (0.00)
ARCH(5)	0.20 [2.27] (0.02)	*	*	*
GARCH(1)	*	0.58 [2.53] (0.01)	1.01 [4.39] (0.00)	-1.77 [-27.58] (0.00)
GARCH(2)	*	0.50 [4.84] (0.00)	-0.12 [-0.87] (0.38)	-2.46 [-25.49] (0.00)
GARCH(3)	*	0.71 [8.47] (0.00)	0.53 [8.22] (0.00)	-1.70 [-19.46] (0.00)
GARCH(4)	*	-0.99 [-4.20] (0.00)	-1.15 [-6.60] (0.00)	-0.90 [-13.33] (0.00)



	ARCH(5)	GARCH(4,0)	TGARCH(5,3)	EGARCH(4,1)
GARCH(5)	*	*	0.50 [2.33] (0.02)	*
RESID(-1)^2*(RESID(-1)<0)	*	*	0.10 [3.36] (0.0008)	*
ABS(RESID(-1)/ @SQRT/GARCH(-1)))	*	*	*	-0.14 [-2.60] (0.009)
(RESID(-1)/ @SQRT/GARCH(-1))	*	*	*	0.02 [0.76] (0.45)
R²	0.12	0.12	0.12	0.11
AIC	-4.67	-4.65	-4.68	-4.69
DW	1.99	1.95	1.97	1.89
Ljung-Box/LB(36) (Prob)	16.001 (0.998)	19.019 (0.991)	17.595 (0.996)	18.932 (0.991)
LB²(36) (Prob)	21.190 (0.976)	33.254 (0.600)	19.310 (0.990)	28.526 (0.808)
Normalité	Oui	Oui	Oui	Oui
Effet ARCH	Non	Oui(lag :5)	Non	Oui(lag :2)

_____ **Commentaire** : Si l'on s'en tient au corrélogramme des résidus aux carrés (Cfr **LB²(36)**), il apparaît que – pour tous les modèles/processus – les résidus issus de la variance conditionnelle sont des bruits blancs (prob>5%).

_____ **Modèle optimal** : au regard des résultats ci-dessus, l'on note le modèle optimal – selon les critères – comme suit :

- o AIC minimum (qualité de l'équation de la moyenne) : **EGARCH(4,1)** ;
- o DW maximal : ARCH(5) ;
- o LB²(36) minimum (qualité de l'équation de la moyenne) : **TGARCH(5,3)** ;
- o Effets ARCH : ARCH(5) et TGARCH(5,3).

Le modèle GARCH(4,0) est éliminé à ce niveau, et quatre modèles sont ainsi candidats à la prévision. Dans le point qui suit, nous nous servons des critères de bonne prévision (Mean Absolute Percentage Error/MAPE minimal et Coefficient d'inégalité de Theil proche de zéro) pour sélectionner le modèle optimal parmi les quatre.

F. Prévision

► Choix du modèle optimal (recours aux critères de bonne prédiction)

Nous recourons aux procédures de simulation dynamique de « EViews » pour prévoir notre série filtrée « DLEURUS » à travers nos quatre modèles (ne disposant pas de nouvelles données, notre série filtrée a été tronquée de 5 mois. Sur base de



cette série tronquée, une ré-estimation nous a aidée à obtenir les prédictions souhaitées).

_____ Pour ce faire, sur EViews, procéder comme suit :

- o Taper : *smpl 1990:01 2012:10*
- o Estimer les paramètres du modèle sélectionné ;
- o après l'estimation des paramètres, dans l'output des résultats, cliquer sur **Forecast** : la boîte de dialogue suivante apparaît (elle complète la procédure. Ex : GARCH(4,0)) :

► Préviation

_____ Les valeurs MAPE et Theil recueillies, pour différents modèles, témoignent de la faiblesse des modèles retenus (aucun modèle n'est performant : maintenir la prédiction sur série brute) : voir tableau ci-dessous.

Critères	Modèles/Processus			
	ARCH(5)	GARCH(4,0)	TGARCH(5,3)	EGARCH(4,1)
MAPE	98.8501	99.8485	98.0892	98.7480
THEIL	0.9546	0.9688	0.9456	0.9523



Références bibliographiques

Antoine Terracol (2008), « *Stata par la pratique : statistiques, graphiques et éléments de programmation* par Eric Cahuzac et Christophe Bontemps », the Stata Journal 8, Number 4, pp. 574-578.

Baltagi Badi H. (2005), « *Econometric Analysis of Panel Data* », 3^e édition, JW edition, Angleterre, 316 p.

Benchimol Jonathan, « *Formation EViews 7 : Introduction* », 95 p.

Bocquier Philippe (1996), « *L'analyse des enquêtes biographiques à l'aide du logiciel STATA* », Documents et Manuels du CEPED n°4, Paris, 224 p.

Bontemps Christophe (2002), « *Stata par la pratique – Partie II : les graphiques de Stata* », 20 p.

Bourbonnais R. (2015), « *Econométrie : cours et exercices corrigés* », 9^e édition, éd. DUNOD, Paris, 392 p.

_____ **(2009)**, « *Logiciel Eviews* », Université de Paris-Dauphine, 31 p.

Bourbonnais, R. et Terraza, M. (2016), « *Analyse des séries temporelles – Applications à l'économie et à la gestion : Cours et exercices corrigés* », éd. Dunod, 4^e édition, Paris, 354 p.

Bozio Antoine (2005), « *Introduction au logiciel STATA* », Paris, 18 p.

Cadore I. et al. (2009), « *Econométrie appliquée : Méthodes – Applications – Corrigés* », éd. de boeck, 2^e édition, Bruxelles, 462 p.

Cadot Olivier (2012), « *Stata pour les nuls* », 65 p.

Casin Philippe (2009), « *Econométrie : Méthodes et applications avec EViews* », éd. Technip, Paris, 224 p.

Charpentier Arthur, « *Cours de séries temporelles : Théories et applications – Volume 2* », 141 p.

Christopher Baum F. (2001), « *Stata : the language of choice for time series analysis* », in The Stata Journal 1, number 1, pp. 1-16.

Couderc Nicolas, « *Econométrie appliquée avec Stata* », Université Paris 1 Panthéon-Sorbonne, 22 p.



Deniu C., Fiori G. et Mathis A. (2015), « Sélection du nombre de retards dans un modèle VAR: Conséquences éventuelles du choix des critères », in *Economie et prévision*, n° 106, 1992-5. Développements récents de la macro-économie, pp. 61-69. (lien : http://www.persee.fr/doc/ecop_0249-4744_1992_num_106_5_5315).

Desjardins Julie (2005), « L'analyse de régression logistique », *Tutorial in Quantitative Methods for Psychology*, vol. 1(1), pp. 35-41.

Doucoure Fodiye B. (2008), « Méthodes économétriques: cours et travaux pratiques », éd. ARIMA, 5^e édition, Dakar, 511 p.

Goaied M. et Sassi S. (2012), « Econométrie de données de Panel sous Stata », 1^{ère} édition, I.H.E.C/LEFA, 45 p.

Gosse, J-B. et Guillaumin, C. (2011), « Christopher A. Sims et la représentation VAR », 15 p.

Hurlin Christophe, « Econométrie des variables qualitatives : Modèles à variable dépendante limitée (Modèles Tobit simples et Tobit Généralisés) », 52 p.

_____ (2003), « Econométrie des variables qualitatives : Modèles Dichotomiques Univariés (Modèles Probit, Logit et Semi-Paramétriques) », 57 p.

_____ (2003), « Econométrie des variables qualitatives : Modèles Multinomiaux (Modèles Logit Multinomiaux Ordonnés et non Ordonnés) », 32 p.

_____, « L'Econométrie des Données de Panel : Modèles Linéaires Simples », 68 p.

_____ (2007), « Modèles ARCH-GARCH: Application à la VaR », Université d'Orléans, 78 p.

Hurlin, C. et Mignon, V. (2006), « Une synthèse des Tests de Cointégration sur Données de Panel », 33 p.

I Gusti Ngurah A. (2009), « Time Series Data Analysis Using Eviews », édition John Wiley and Sons, 635 p.

Kenneth Simons L. (2013), « Useful Stata Commands (for Stata version 12) », 47 p.

Kintambu Mafuku E.G. (2004), « Principes d'Econométrie », Presses de l'Université Kongo, 4^e édition, 285 p.

Kpodar Kangni (2007), « Manuel d'initiation à Stata (version 8) », CERDI, Clermont-Ferrand, 97 p.



Lubrano Michel (2008), « Modélisation Multivariée et Cointégration », 32 p.

_____ (2008), « Tests de Racine Unitaire », 46 p.

Luyinduladio Menga E. (2009), « Manuel d'initiation à EViews », inédit, Septembre, 79 p.

Michée Sendula, « Guide d'utilisation Stata 9 », inédit, 44p.

Mignon Valérie (2008), « Econométrie: Théorie et Applications », éd. ECONOMICA, 236 p.

Nicholas L., Hébert B-P et Laplante B. (2007), « introduction à Stata », 46 p.

Ouellet Estelle (2005), « Guide d'économétrie appliquée pour Stata pour ECN 3950 et FAS 3900 », Université de Montréal.

Park Hun Myoung (2008), « Univariate Analysis and Normality Test Using SAS, Stata, and SPSS », Working Paper, The University Information Technology Services (UITs) Center for Statistical and Mathematical Computing, Indiana University, 41 p.

Pellier K. (2007), « Travaux Dirigés d'Econométrie – M1: Guide d'utilisation d'EViews », 14 p.

Pétry F. et Gélinau F., « Guide pratique d'introduction à la régression en sciences sociales – Deuxième édition revue et augmentée », 42 p.

Quantitative Micro Software/QMS (2004), « EViews 7 Command and Programming Reference », USA, Avril, 580 p.

Quantitative Micro Software/QMS (2004), « EViews 7 Object Reference », USA, Novembre, 764 p.

Rabe-Hesketh S. et Everitt B.S. (2004), « A Handbook of Statistical Analyses using Stata », 3è édition, éd. CHAPMAN and HALL/CRC, Londres, 304 p.

Ricardo Perez-Truglia (2009), « Applied Econometrics using Stata », Havard University, 170 p.

Robert Alan Y. (2007), « Robust Regression Modeling with STATA – lecture notes », 93 p.

_____ (2007), « Stata 10 (Time Series and Forecasting) », in Journal of Statistical Software, volume 23, Software Review 1, december, 18 p.

Robert de Jong (2003), « Eviews mini manual », 6 p.



Robert Dixon, « *Simulation of the Klein–Goldberger Model using Eviews* », 6 p.

Rous Philippe, « *Modèles Estimés sur Données de Panel* », cours d'Econométrie des données de panel – Master « Economie et Finance », Université de Limoges, 76 p.

Tombola Muke C. (2012), « *Séminaire d'Economie Mathématique I-avec initiation aux logiciels EIEWS, STATA et MATLAB/modules 1 & 2* », Laréq, Décembre, 37 p. (www.lareq.com).

Tsasa Vangu K. JP. (2012), « *Introduction à la programmation à l'aide du logiciel EViews – avec 11 programmes exécutables sur EViews : illustrations + commentaires* », Laréq, Décembre, 19 p. (www.lareq.com).

Vescovo Aude, « *Cours sur le logiciel Stata* », IRD-AFRISTAT, 40 p.

